

# Chapter 10

## Different Views on the Finger— Score-Level Fusion in Multi-Perspective Finger Vein Recognition



Bernhard Prommegger, Christof Kauba and Andreas Uhl

**Abstract** In finger vein recognition, the palmar view of the finger is used almost exclusively, with some exceptions where the dorsal view is utilised. Only little attention has been paid to all other views around the finger's longitudinal axis. We established a multi-perspective finger vein dataset comprising of views all around the finger's longitudinal axis, captured using our self-developed rotating multi-perspective finger vein capture device. The performance of the single views is evaluated using common finger vein recognition algorithms. Based on these single view scores, several score-level fusion experiments involving different fusion strategies are carried out in order to determine the best performing set of views and feature extraction methods to be fused in terms of recognition accuracy while minimising the number of views involved. Our experimental results show that the recognition performance can be significantly improved over the best performing single view one with as few as two views and two-feature extraction methods involved.

**Keywords** Finger vein recognition · Multi-perspective fusion · Biometric fusion · Score-level fusion · Multi-algorithm fusion · Multi-perspective finger vein capture device · Finger vein dataset

---

B. Prommegger (✉) · C. Kauba · A. Uhl  
Department of Computer Sciences, University of Salzburg,  
Jakob-Haringer-Str. 2, 5020 Salzburg, Austria  
e-mail: [bprommeg@cs.sbg.ac.at](mailto:bprommeg@cs.sbg.ac.at)

C. Kauba  
e-mail: [ckauba@cs.sbg.ac.at](mailto:ckauba@cs.sbg.ac.at)

A. Uhl  
e-mail: [uhl@cs.sbg.ac.at](mailto:uhl@cs.sbg.ac.at)

© The Author(s) 2020

A. Uhl et al. (eds.), *Handbook of Vascular Biometrics*, Advances in Computer Vision and Pattern Recognition, [https://doi.org/10.1007/978-3-030-27731-4\\_10](https://doi.org/10.1007/978-3-030-27731-4_10)

## 10.1 Introduction

Finger vein recognition as one representative of vascular pattern biometrics deals with the vascular pattern inside the fingers of a human. Since one of the first mentions of finger veins as a biometric trait in academia by Kono [1] in 2000, they have received much attention not only from academia but also from industry. Commercial off-the-shelf (COTS) finger vein capture devices, as well as most research papers solely, use the palmar (front side of the finger) view in combination with light transmission (the light source and the image sensor are placed on opposite sides of the finger) as illumination source. Multi-perspective finger vein recognition deals with two or more arbitrary perspectives around the finger's longitudinal axis. Despite the advantages of multi-perspective finger vein biometrics over single view ones, these additional perspectives have not got much attention so far. Moreover, there is no publicly available multi-perspective finger vein dataset yet.

This chapter is based on our previous work [2] where we designed a novel, multi-perspective finger vein capture device in order to establish the first multi-perspective finger vein data set. This dataset comprises of images captured all around the finger's longitudinal axis in  $1^\circ$  steps. Based on this dataset, each of the different views has been evaluated individually and some simple fusion experiments have been conducted. The main focus of this chapter is on the fusion of multiple perspectives and feature extraction methods in order to determine the best performing combination in terms of recognition accuracy by employing a more advanced multi-sample score-level fusion scheme as well as by applying further fusion strategies in terms of view and feature combinations. We analyse all possible pairs and triples of perspectives and all possible combinations of the used feature extraction methods. In addition, we combine the best results of our multi-perspective and multi-algorithm fusion experiments to one single combined fusion. Our main goal is to minimise the number of views and feature extraction methods involved, while maximising the recognition accuracy. A typical multi-perspective finger vein capture device contains one image sensor and one light source situated at the right position per desired view. The more views are to be captured, the more camera and illumination modules have to be equipped, thus increasing the production costs, the complexity and the overall size of the finger vein capture device. If the number of desired perspectives is further increased, the construction of a suitable capture device is no longer feasible without the need of rotating parts. Our current multi-perspective finger vein capture device is such a rotating device, making it more susceptible to malfunctions and external influences than a capture device containing no rotating parts. Moreover, the capturing time is increased as the capture device has to rotate all around the finger. Hence, it is beneficial to reduce the number of different views to be captured to a minimum in order to reduce the complexity and production costs of the biometric capture device and to avoid the need for a rotating device while still preserving the advantages of a multi-perspective capture device.

The rest of this chapter is structured as follows: Sect. 10.2 starts with a description of multi-perspective finger vein biometrics including related work regarding other

views than the palmar and dorsal one in finger vein recognition. Our multi-perspective finger vein capture device design is described in Sect. 10.3. Section 10.4 introduces our multi-perspective finger vein dataset captured with the aforementioned device. Section 10.5 gives an overview of biometric fusion in general followed by related work on biometric fusion in finger vein recognition. Section 10.6 explains our experimental set-up, including the finger vein recognition tool chain as well as the fusion framework we utilised and lists the experimental results, followed by a results discussion. Section 10.7 concludes this paper and gives an outlook on future work.

## 10.2 Multi-perspective Finger Vein Biometrics

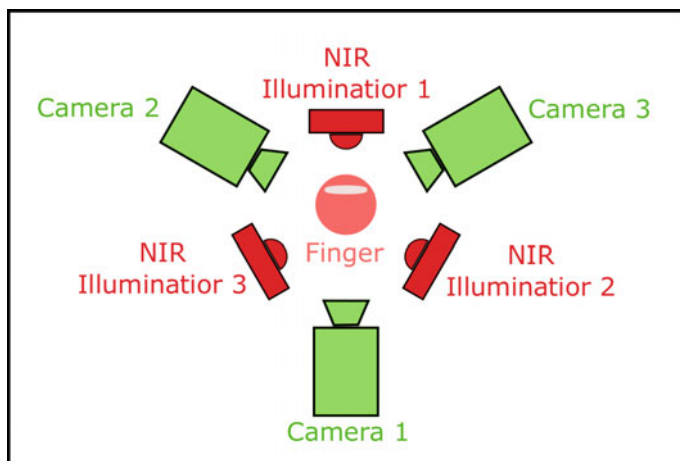
The majority of the available finger vein recognition schemes as well as all available COTS finger vein capture devices deal with the palmar (also called ventral) view of the finger. There are only some exceptions where the dorsal view is used. Raghavendra and Busch [3] proposed the first dorsal finger vein acquisition and a complete recognition tool chain including several different feature extraction schemes. In the scope of the PROTECT project (<http://www.projectprotect.eu>), we acquired the first publicly available dorsal finger vein dataset [4] using the predecessor of our open-source finger vein capture device. In [5], we established a larger dorsal finger vein dataset captured using both of our proposed open-source finger vein capture devices, which design is described in Chap. 3 of this book [6].

There are more views around the finger than the palmar and dorsal one that can be captured. A single finger is an elliptical cylinder-shaped object, hence, there are all possible views around its longitudinal axis ( $360^\circ$  of rotation) available. Multi-perspective finger vein recognition describes the use of two or more of these perspectives around the finger's longitudinal axis. Multi-perspective finger vein recognition has several advantages over the single perspective one: The vein patterns of the palmar and dorsal view as well as of the perpendicular views are independent from each other [7]. By fusing more than one perspective that is independent enough from each other (i.e. the rotation angle between the single perspectives has to differ enough for the perspectives to be independent of each other), the overall recognition performance can be increased easily. Tome et al. [8, 9] showed that finger vein and hand vein recognition systems are susceptible to a simple type of presentation attack. By using a paper printout of the vein pattern, they were able to successfully spoof several finger vein capture devices. This paper printout is a flat, 2D representation of the vein pattern. If a biometric capture device takes finger vein images from different perspectives, such simple 2D printout attack finger vein presentation will not be identified as bona fide finger vein presentation. Thus, a multi-perspective finger vein capture device is successfully able to prevent this kind of presentation attack. However, multi-perspective finger vein recognition bears some disadvantages too: The biometric capture devices get more complex, either more than one camera and illumination module are needed, or the capture device has to be built in a rotating manner. This leads to higher production costs of multi-perspective capture devices

and especially rotating capture devices are more error prone due to the moving parts. Another disadvantage is the bigger size of a multi-perspective capture device compared to single perspective ones. The multiple image sensors/illuminator modules or the rotating parts need more space than just a single image sensor in combination with one illumination module.

Lu et al. [10] proposed a multi-perspective finger vein recognition system using two cameras. The cameras are placed at an angle of  $60^\circ$  next to each other, each camera is located  $30^\circ$  apart from the palmar view. They applied feature—as well as score-level fusion using the two views captured simultaneously by the two cameras and were able to improve the recognition performance of the single view ones. Zhang et al. [11] employed a binocular stereoscopic vision device to do 3D point cloud matching of hand veins and knuckle shape. Their capture device set-up consist of two cameras, placed in a relative position of about  $45^\circ$  next to each other, each one equipped with an NIR-pass filter. There is only a single light transmission illuminator placed underneath the palm of the hand. The 3D point clouds are generated by extracting information from the edges of the hand veins and knuckle shapes and then compared utilising a kernel correlation method, especially designed for unstructured 3D point clouds. The authors claim that their proposed method is faster and more accurate compared to 2D vein recognition schemes. In [12] the authors propose a 3D hand vein capturing system based on a rotating platform and a fixed NIR camera. The camera is located above the hand, the hand is put on a handle with an integrated light transmission illuminator. This handle is mounted on a rotating plate. Then the plate rotates around the z-axis. However, the degree of rotation is limited due to the limited movement of the hand in this position. A 3D point cloud is generated from the single view images and matched using kernel correlation. This should help to overcome hand registration and posture change problems present in hand vein recognition if only 2D vein patterns/images are available.

Nevertheless, true multi-perspective finger vein recognition (evaluating more than two different views around the finger) has not been investigated so far, except for our previous work [2]. One reason herefore might be the lack of available multi-perspective finger vein datasets. In order to acquire such a dataset a suitable biometric capture device, able to capture the different views to be acquired, is essential. Capturing these additional perspectives could be done by utilising either a COTS capture device or one of the capture devices proposed in other works by simply turning the finger around its longitudinal axis. However, it is difficult to position the finger in the correct rotational angle. Thus, rotating the finger itself implies the disadvantage of an inaccurate rotation angle and deviations in the rotation angle across different iterations, leading to a low repeatability and a low quality dataset. In order to acquire a suitable multi-perspective finger vein dataset comprising of images captured in several, defined perspectives, either a biometric capture device comprising of several cameras and illumination modules, able to capture more than one view simultaneously, or a rotating biometric capture device able to capture these views consecutively, is necessary. If only a limited number of perspectives are involved, a suitable biometric capture device can be built without any rotating parts, just by equipping an individual image sensor and an associated illumination module per desired



**Fig. 10.1** Multi-perspective finger vein set-up exhibiting three different perspectives based on three image sensors and three illuminator modules

view (an example with three different views is shown in Fig. 10.1). The illumination intensity has to be adjusted per view as the path to penetrate the finger is different for each individual view, requiring a stronger or weaker illumination depending on the distance. If more perspectives are desired, rotating the capture device around the finger while the finger remains in a fixed position during the acquisition process is the only feasible option.

The design and construction of a practicable biometric capture device is a complex task. Furthermore, the actual data acquisition is a tedious and time-consuming work. In our previous paper [2], we proposed a rotating multi-perspective finger vein capture device that is able to capture the finger all around its longitudinal axis ( $360^\circ$ ). We established a multi-perspective finger vein dataset consisting of 252 individual fingers. Based on this dataset, we evaluated the different views around the finger in  $5^\circ$  steps and concluded that the palmar followed by the dorsal one achieve the best single view recognition performance. Moreover, we applied a simple score-level fusion strategy and showed that the recognition performance can be improved by fusing more than one view. This chapter is an extension of our previous work. Based on our proposed multi-perspective finger vein capture device, we refine and extend our previous results by the following:

- Improving the recognition tool chain to improve the single view results, especially the ROI extraction and by including a new recognition scheme proposed by Matsuda et al. [13].
- Employing an advanced score-level fusion framework (BOSARIS [14]).
- Exploring different fusion strategies in terms of which views to include in the fusion.
- Evaluating multi-algorithm fusion per view (fusion is done at score level).

- Combining multi-perspective and multi-algorithm fusion.

The purpose of our evaluations is to maximise the recognition performance while minimising the number of single views involved. If only a limited number of views is involved, the capture device can be built without the need for any rotating parts just by equipping an individual image sensors and an illumination modules per desired view. A biometric capture device which relies on rotating parts is more error prone and more susceptible to external influences, the rotation speed can vary due to increased friction or it can be completely blocked if the finger is not properly inserted. The rotating parts exhibit a higher wear than non-moving parts and are thus more prone to failures. Moreover, the acquisition time of a rotating capture device is higher compared to a non-rotating one as the device needs to rotate around the finger in order to capture the different views. Furthermore, a capturing device exhibiting a closed box design, where the capture subject has to put his finger into a “black hole” poses psychological disadvantages and leads to discomfort. Hence, in practical applications of multi-perspective finger vein biometrics only a capture device built in a non-rotating and open manner is feasible. Consequently, we aim to identify the best combination of two or three views to include in the fusion in order to build such a multi-perspective finger vein capture device based on fixed, non-moving parts only. Figure 10.1 shows the schematic principle of such a capture device for three perspectives: it consists of three independent image capturing pairs, each consisting of its own NIR illumination module and NIR camera.

### 10.3 Multi-perspective Finger Vein Capture Device

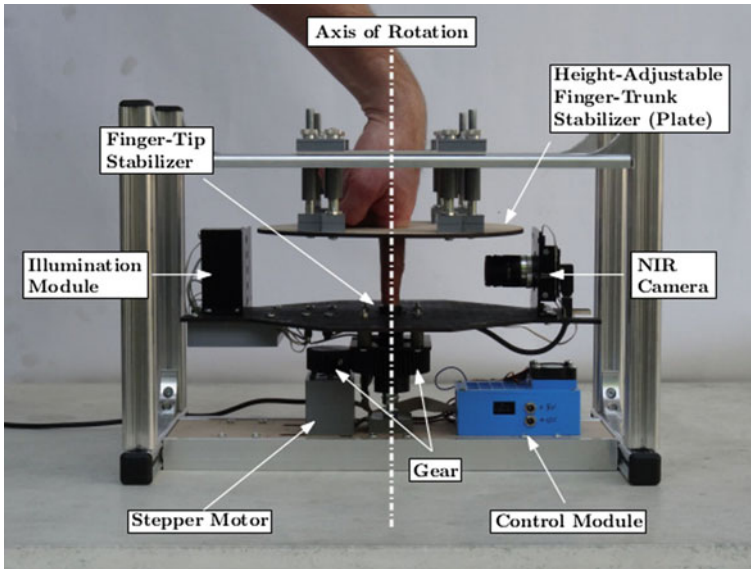
In order to acquire a multi-perspective finger vein dataset, we designed a custom finger vein capture device tailored to this purpose. For more details on the general principle of a finger vein scanner and the vascular pattern recognition basics, the interested reader is referred to our open finger vein scanner chapter [6] and the introductory chapter [15] of this book, respectively. Our multi-perspective finger vein capture device is able to capture images from all around the finger’s longitudinal axis ( $360^\circ$ ). An illustration of the unwrapped finger vein capture device with all its parts labelled can be seen in Fig. 10.2. Its outside dimensions (of the aluminium frame including the rotating part) are  $258 \times 325 \times 455$  mm (width  $\times$  height  $\times$  depth). The rotating part (rotator) has a diameter of 380 mm. The device consists of an aluminium frame, where the rotation motor and the control board are located and a rotator, which rotates around the finger. The rotating part is connected to a stepping motor by two cogwheels. These cogwheels have a gear ratio of 1:5/3 (motor to rotor). The stepping motor (SY42STH47-1684A [16]) which drives the rotator has 200 steps per full rotation ( $1.8^\circ$  per single step). We use a micro-stepping of 1/16, thus one step corresponds to  $0.0675^\circ$ . Hence, it is possible to capture a maximum of 5333 different perspectives of the finger. Located on the right side of the device is the image sensor, an IDS Imaging UI-1240ML-NIR industrial NIR-enhanced camera

[17]. It has a max. resolution of  $1280 \times 1024$  pixels, a max. frame rate of 25 fps and is equipped with a Fujifilm HF9HA-1b 9mm 2/3" wide-angle lens [18]. To reduce the influence of ambient light, an additional NIR long-pass filter (MIDOPT LP780 [19], with a cut-off wavelength of about 750 nm and a useful range of 780–1000 nm) is mounted on top of the camera lens. The illumination module is located on the opposite side of the image sensor (the left side in Fig. 10.2). Our multi-perspective finger vein capture device is based on the light transmission principle. Instead of typical NIR LEDs the illumination module consists of five NIR laser modules with a peak emission wavelength of 808 nm placed in a strip. Laser diodes have several advantages over LEDs, especially, if the finger is not placed directly on top of the illumination module as mentioned in Chapter [6]. Due to the rotating principle of the biometric capture device, it is not possible for the finger to touch the illumination module, which prevents the use of LEDs without impacting the image quality. Each laser module consists of a NIR laser diode, a control PCB for the laser diode and a housing with a focus-adjustable lens. The plane of focus of the laser modules is set at the axis of rotation where the finger is placed, leading to the highest possible amount of illumination at the position of the finger. Each of the laser modules can be brightness controlled separately (by adjusting the operating current) and independently, enabling a uniform illumination along the whole finger. The finger is put into the capture device at its axis of rotation (in the centre of the image in Fig. 10.2). A fingertip stabiliser (a custom 3D printed part which inside is shaped like the outside of a fingertip) is located at the inside bottom of the rotating part and a height-adjustable finger trunk stabiliser, which is basically a wooden plate with a hole in the middle is located above the rotating part. These finger stabilisers help to reduce finger movements during one acquisition run to a minimum. The finger is put into the capture device so that its tip is inside the fingertip stabiliser, pushing the height-adjustable plate down. Afterwards, this individual finger height is fixed using four screws on the top of the scanner and remains fixed until a new finger is to be captured. All parts except the stepping motor, the camera including the lens and NIR long-pass filter) are self-designed and manufactured by ourselves, including several 3D printed parts, the wooden housing of the rotating part, the housing of the control board, the control board itself and the aluminium frame.

The acquisition process is semi-automated. At first, the subject has to put the finger into the device. Then the height of the finger trunk stabiliser plate has to be adjusted and the operator initiates one capturing run ( $360^\circ$  around the finger's longitudinal axis), starting the automated part of the acquisition process.

During this automated data acquisition part, the illumination for each laser module is set automatically by the help of an automated brightness control algorithm. This algorithm tries to achieve a sufficient and uniform illumination along the finger in order to obtain an optimal image contrast. It evaluates the average grey level of the image area around the centre of each laser module  $i$  ( $GL_{current}^i$ ) and compares this value to a predefined target grey level ( $GL_{target}^i$ ). If there is a deviation between these two values, the operating current of the corresponding laser module is adjusted:

$I_{corr}^i = \frac{GL_{target}^i - GL_{current}^i}{GL_{max}} \cdot \frac{I_{max}}{2 \cdot n}$ , where  $GL_{max}$  is the maximum grey value (255 for 8 bit



**Fig. 10.2** Self-designed multi-perspective finger vein capture device (image originally published in [2], ©2018 IEEE)

images) and  $n$  is the number of the current iteration. Initially, all laser modules are set to half of their maximum operating current  $I_{\max}$  (corresponding to its maximum intensity). The algorithm finishes in at most  $\log_2(I_{\max})$  steps.

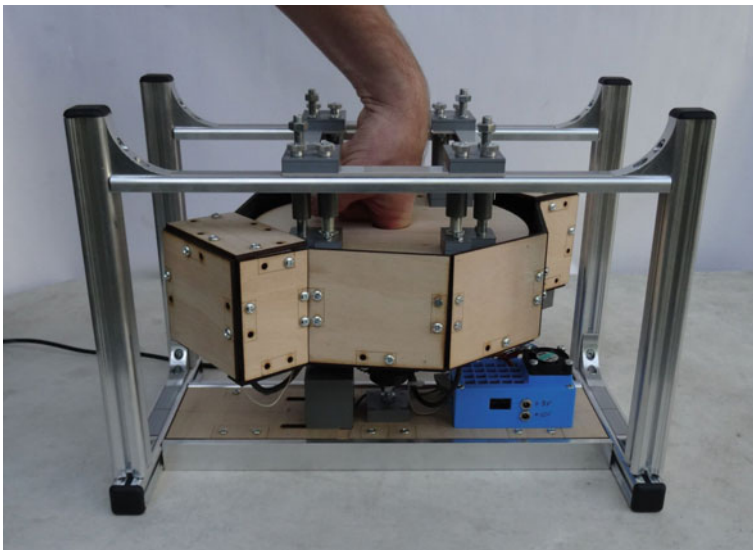
After the optimal intensity level for each laser module is set, the video sequence recording is started. The rotator starts to rotate around the finger and an indicator LED is turned on to synchronise the video stream. The rotation is stopped when the rotator reaches its start position again and at this point the indicator LED is turned off. A few frames later the video sequence recording is stopped too. The videos are recorded in the MP4 container format using the MJPG video codec with a frame rate of 15 fps and YUV colour space. The speed of the rotation and the video frame rate are synchronised such that a defined resolution (in degree) of images per full rotation (video frames) is met and the desired degree steps can later be extracted from single, individual frames without the need for temporal interpolation. The set illumination intensity remains the same for the whole capturing run until all perspectives are captured. This ensures the compatibility and comparability of the single, individual perspectives to each other. The different projections in  $1^\circ$  steps corresponding to single video frames are then extracted out of the video sequence. The capture device's indicator LED is utilised to synchronise the video frames with the beginning and the end of the rotation. In theory, there should be 361 images per full rotation run ( $0^\circ$  and  $360^\circ$  is captured separately). Due to slight variations in the rotation speed and the video frame rate, there are between 357 and 362 frames instead of 361. Thus, it



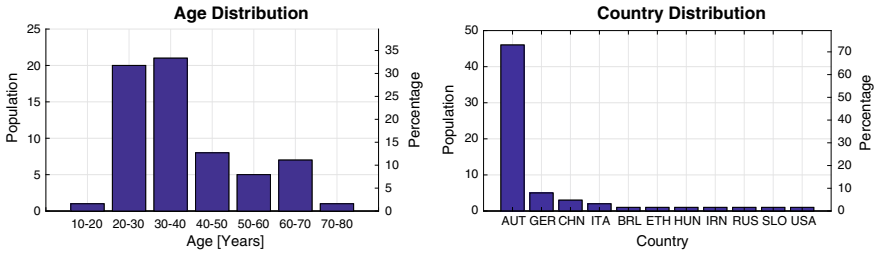
became necessary to map the frame with the minimum deviation from the desired rotational angle to the corresponding perspective, resulting in a maximum deviation of  $0.5^\circ$  from the desired rotation angle.

## 10.4 Multi-perspective Finger Vein Dataset

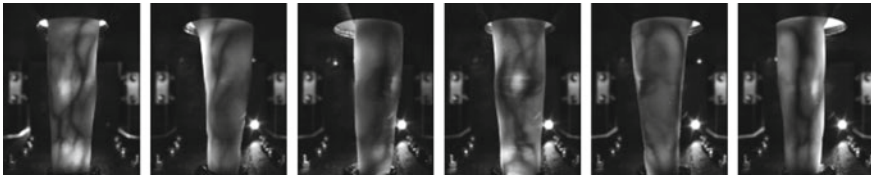
With the help of our self-designed multi-perspective finger vein capture device, we established a multi-perspective finger vein dataset in order to be able to conduct our multi-perspective score-level fusion experiments. This dataset currently consists of 63 subjects, 4 fingers per subject (index and middle finger of the left and right hand) and 5 runs per finger. The thumb and the pinky finger were not included as they are too short compared to the index and middle. The ring finger was skipped as well as it turned out to be too uncomfortable for the subjects to put it in the capture device for the whole capturing process. The finger was removed and inserted in the device again after each run. During each run, a video sequence of a full  $360^\circ$  rotation with a target resolution of  $1^\circ$  (each frame corresponds to a  $1^\circ$  step) is captured. Figure 10.3 shows the capture device during the data acquisition process. The acquisition process takes approximately 45 s per capture attempt, hence it takes about 15 min to capture a single subject, including all four fingers, 5 runs per finger. The whole dataset consists of  $63 \times 4 \times 5 \times 361 = 454,860$  images in total. The extracted video frames have a resolution of  $1024 \times 1280$  pixels and are 8-bit greyscale images stored in png format.



**Fig. 10.3** Data acquisition with the multi-perspective finger vein capture device (image originally published in [2], ©2018 IEEE)



**Fig. 10.4** Age (left, image originally published in [2], ©2018 IEEE) and country of origin distribution (right) for the multi-perspective finger vein dataset



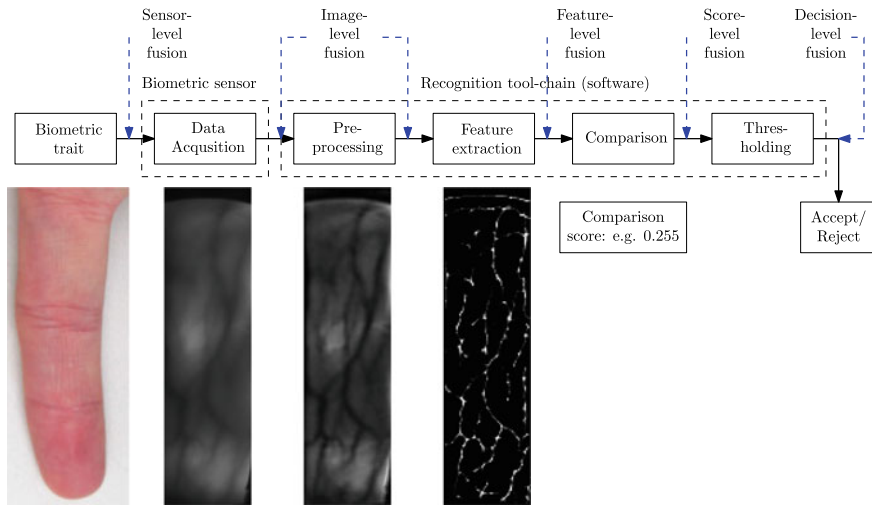
**Fig. 10.5** Multi-perspective finger vein dataset example images, from left to right: 0°, 60°, 120°, 180°, 240°, 300° (image originally published in [2], ©2018 IEEE)

The finger is always located in the centre area of the image, thus the images are then cropped to  $650 \times 1280$  pixels to retain the usable finger area only. Figure 10.5 shows some example images in different perspectives from 0° to 300°. It can be clearly seen that the visible vein lines vary among the different perspectives. The black part at the centre top area in the images results from the finger trunk stabilisation plate, which is pushed in further or less depending on the length of the finger.

The gender distribution of the 63 subjects is almost balanced with 27 (42.7%) female and 36 (57.3%) male subjects. The subjects represent a good cross section among all different age groups, as the age distribution, depicted in Fig. 10.4 left, shows. There is only a slight overhang among the 20–40 year old subjects. The youngest subject was 18 and the oldest one 79 years old. The subjects are from 11 different countries (Austria, Brazil, China, Ethiopia, Hungary, Iran, Italy, Russia, Slovenia, USA) while the majority of subjects are white Europeans (73%). The origin country distribution is depicted in Fig. 10.4 right. The dataset is available for research purposes and can be downloaded at <http://wavelab.at/sources/PLUSVein-FR/>.

## 10.5 Biometric Fusion

Like every typical biometric recognition system, a finger vein recognition system consists of five steps/modules: image acquisition, preprocessing, feature extraction, comparison and the final decision. This recognition tool chain is depicted in Fig. 10.6.



**Fig. 10.6** Basic components of a biometric recognition system including the different levels of fusion by taking the example of finger veins (second row)

There are two modes, enrolment and authentication. Authentication includes both, verification as well as identification. During enrolment one or several finger vein images are captured and the extracted biometric templates are stored in a database. During authentication a new template is extracted from a newly captured image and compared against one or more templates stored in the database. The result is a comparison score. Finally the decision module outputs for the capture subject an “accept” or “reject” depending on the evaluation of the comparison score against a threshold.

According to the ISO/IEC TR 24722:2015 standard [20], biometric fusion can be regarded as a combination of information from multiple sources, i.e. sensors, characteristic types, algorithms, instances or presentations in order to improve the overall system’s performance and to increase the systems robustness.<sup>1</sup> Biometric fusion can be categorised according to the level of fusion and the origin of input data. The different levels of fusion correspond to the components of a biometric recognition system:

- **Sensor-level fusion:** is also called multisensorial fusion and describes using multiple sensors for capturing samples of one biometric instance [20]. This can either be done by the sensor itself or during the biometric processing chain. An example of sensor-level fusion are finger vein images that have been captured using different wavelength of near-infrared light and fused by merging the different wavelength bands to obtain one single output image. This can be done by a single biomet-

<sup>1</sup>Recognition performance is just one aspect. PAD performance (robustness against presentation attacks) is another aspect to keep in mind.

ric capture device. Another example is the acquisition and fusion of fingerprint images captured using optical, electrostatic and acoustic sensors.

- Image-level fusion: during data acquisition, the biometric capture device itself might be able to capture multiple samples of the same biometric trait and combine those samples to a single output sample. Image-level fusion corresponds to fusing several images captured from the same biometric trait but not necessarily within the sensor device. Image-level fusion can also be applied after preprocessing so the input to the fusion module is the preprocessed images. One example of image-level fusion is a finger vein capture device that captures more than one finger simultaneously and combines the images from the individual fingers into a single output image, which is also called multi-instance.
- Feature-level fusion: during template creation, several meaningful features, describing the biometric trait's properties, are extracted from the preprocessed images and stored in a feature vector, commonly denoted as biometric template. Feature-level fusion combines several such feature vectors to form a new, higher dimensional feature vector which should represent a subject's biometric traits in a different and more discriminant way. Dimensionality reduction methods are beneficial in combination with feature-level fusion to extract the most significant and discriminative features and to save storage space.
- Score-level fusion: during the comparison step, two templates are compared against each other and a similarity or dissimilarity score is calculated. Score-level fusion combines two or more of those scores into a new, single score. The input scores can originate from different comparison modules. They should either be compatible with each other (e.g. all are similarity scores exhibiting the same range of possible values) or else a score normalisation technique has to be applied during the fusion.
- Decision-level fusion: the output of the decision module is a binary one, which can be interpreted as match/non-match or accept/reject. Decision-level fusion combines two or more of these binary output decisions to a single output one. Usually, majority of voting schemes are employed at decision-level fusion. Note that at the decision level, the least information is available (only a binary decision), compared to the other levels of fusion.

Regarding the origin of the input data, biometric fusion can be categorised into:

- Multi-modal fusion: multiple different types of biometric traits from the same subject is fused together. A popular example is the fusion of information from fingerprints and finger veins or iris and periocular.
- Multi-instance fusion: multiple instances of the same type of biometric trait are fused together. For example, several finger vein images from different fingers of the same subject or information from both irises of one subject are fused together.
- Multi-presentation fusion: multiple samples of the same instance of biometric trait is captured and fused, e.g. several finger veins of the same finger is captured and fused together.

- **Multi-algorithmic fusion:** multiple feature representations are generated using the same input data, e.g. several different finger vein features are extracted with different algorithms from the same input image and fused together.

There is no direct dependency between the origin of the input data and the level of fusion that is employed.

### ***10.5.1 Fusion in Finger Vein Recognition***

This subsection provides an overview of related work in biometric fusion involving finger veins. The first subsection discusses several single modality fusion approaches. The second subsection lists multi-modality fusion approaches which include finger veins among other biometric traits.

#### **10.5.1.1 Single Modality (Finger Vein Only) Fusion**

Table 10.1 gives an overview of related work on single modality fusion in finger vein recognition, i.e. only data from finger veins is utilised during fusion at different levels. The table lists the level of fusion applied, the origin of the input data to the fusion, the number of images and subjects contained in the used dataset, the reported biometric performance (EER if not stated otherwise) and the year of publication, sorted according to fusion level and year of publication. All the related works listed in Table 10.1 are described in the following.

Yang and Jia [21] presented a multispectral finger vein fusion approach by fusing enhanced finger vein images captured in different wavelengths. They applied an image denoising method followed by image registration and a brightness adjustment prior to the image-level fusion of images captured in six different wavelength bands. Their image-level fusion strategy operates pixel-wise and is based on an improved regional energy integration method in the spatial domain. The comparison scores are obtained by phase-only correlation. They achieved a minimum EER of 11.02% by fusing all six bands.

Guan et al. [22] applied feature-level fusion to Wavelet transform based vein image features. The high- and low-frequency Wavelet features are obtained independently and then fused by a simple nearest-neighbour rule. They did several experiments using different training set sizes and arrived at a maximum recognition rate of 94.35%. Yang and Zhang [23] proposed a feature-level scheme using global and local features. The local features are extracted using a Gabor filter framework and the global ones using 2D invariant moments. The fusion itself is performed by a weighted fusion strategy based on canonical correlation analysis. They reported a lowest FAR of 1.15% and a FRR of 2.47% for their fused features. Gupta and Gupta [24] proposed a feature-level fusion approach of two distinct binary vein features (the features are binary vein images). The first type of features is extracted using repeated

**Table 10.1** Related work in single modality finger vein fusion, ordered according to fusion level and year of publication

Reference	Fusion level	Origin	Images/subjects	Performance (EER)	Year	
[21]	Image	Multi-sample	5760/60	11.02%	2012	
[22]	Feature	Single-sample	2044/292 (fingers)	Recognition rate: 94.35%	2009	
[23]		Single-sample	640/64	FAR: 1.15%, FRR: 2.47%	2010	
[24]		Single-sample	3132/156	2.98%	2015	
[26]		Single-sample	1440/60	0.19%	2016	
[27]		Score	Single-sample	1200/100	0.28%	2010
[28]	Multi-instance		1440/80	0.83% (fusion of 3 fingers)	2012	
[29]	Single-sample		4000/50	0.011%	2012	
[30]	Single-sample		4080/30	1.56%	2013	
[31]	Single-sample		4260/71 (680/85)	2.63%/0.78%	2013	
[32]	Single-sample		3804/634 (fingers)	2.84%	2013	
[33]	Single-sample		1440/60	0.27%	2014	
[2]	Multi-sample		454860/63	0.04%	2018	
[35]	Decision		Single-sample	1620/54	FAR: 0.0086% at 1% FRR	2009

line tracking [25]. The second type of features is obtained by multi-scale matched filtering. A variational approach is proposed to fuse both feature extraction methods. The score calculation is conducted by first aligning the two input images with the help of an affine transformation. The affine transformation matrix is found using a gradient descent optimisation based on a sum of squared differences cost function. The authors report a minimum EER of 2.98%. Kauba et al. [26] used different binary vein feature extraction schemes and applied several advanced feature-level fusion schemes (COLLATE, STAPLE, STAPLER), which were originally proposed for segmentation of magnetic resonance imaging (MRI) brain images together with simple average and majority voting based fusion in the finger vein domain. They conducted two different sets of experiments exhibiting two different fusion strategies. In the first one, only a single feature extraction scheme was used with a set of several different feature extraction parameters per input image. The output features

obtained for the individual parameters where then fused together. In the second set, different feature extraction schemes were applied per input image and their outputs were fused. The authors showed that both strategies (single feature extractor as well as multiple feature extractors) lead to an improvement in the recognition accuracy. The best EER achieved for the first strategy was 0.29% and for the second one 0.19% compared to the best EER for the single features of 0.47%.

Zhou and Kumar [27] proposed a score-level fusion scheme for palm vein recognition based on multiple representations. They extracted four different kinds of features, two based on their proposed representations. The first ones are using Hessian phase information from the vein images, the second ones using localised Radon transform to generate a kind of orientation encoding. The other two ones are based on Ordinal Code and a Laplacian representation, respectively. These four feature representations are compared individually to get the output scores which are then fused by applying a heuristic fusion rule. The authors arrived at a minimum EER of 0.28%. Yang et al. [28] did a score-level fusion of extracted features from multiple fingers of the same subject. They used LBP based features and a Hamming distance based comparison module to generate the scores. These scores are then fused using a simple sum rule in combination with triangular norm. Their best reported EER of 0.83% was achieved by fusion ring, middle and index finger using Frank's t-norm. In [29] Kang Park used local as well as global vein features in combination with score-level fusion. The local features are extracted by the help of LBP and compared using the Hamming distance. The global ones are Wavelet transform based features which are compared using the Euclidean distance. The comparison scores are then fused with the help of a radial basis function based support vector machine. Park reported a best achieved EER of 0.0011%. Liu et al. [30] proposed a score-level fusion scheme including pixel as well as super-pixel based finger vein features. LBP, vein pattern structure based and vein minutiae based features form the pixel based features. The super-pixel based image segmentation is done using the SLIC method. Histogram, gradient and entropy features extracted from the super-pixel based segmentation are then combined and form the super-pixel based features. An Euclidean distance based comparison of both individual features is performed to calculate the comparison scores. These scores are normalised and fused by using the weighted average fusion strategy. The weights are tuned to achieve an optimal EER. They reported a minimum EER of 1.56%. Qin et al. [31] applied score-level fusion to multiple representations of the same finger vein pattern. The vein pattern is represented by three different types of features: finger vein shape based, finger vein orientation based and SIFT feature point based features. The former two are subregion partitioned and subregion compared with the help of the SIFT based features, which are treated individually, leading to three comparison scores. The scores are normalised using the Z-score normalisation and then fused by applying a weighted-sum rule based fusion as well as a support vector machine based fusion. They achieved minimum EERs of 2.63 and 0.78%. Lu et al. [32] proposed a score-level fusion scheme based on Gabor features. Usually, the individual filter responses obtained from the Gabor filter bank are weighted and/or directly combined into a single output feature. Instead, the authors extract and compare the output of each single Gabor filter channel separately.

The corresponding comparison scores are then fused using a simple weighted-sum rule. The authors were able to get an EER of 2.84% using their proposed method. Kauba et al. [33] tested different preprocessing cascades in order to improve the individual performance of the single finger vein feature extraction schemes. Binary and SIFT/SURF based features were compared individually to obtain the output scores. These scores were normalised using Min-Max normalisation and then fused using weighted sum/product/average/minimum/maximum fusion rule. The best fusion rule in terms of lowest EER was chosen accordingly. They were able to achieve a minimum EER of 0.27% with the help of score-level fusion compared to a minimum EER of 0.47% for the single features. In our previous work [2], we performed a multi-sample score-level fusion of several different perspectives around the finger. Therefore, we established a multi-perspective finger vein dataset with the help of our self-designed multi-perspective finger vein capture device, described in Sects. 10.4 and 10.3, respectively. Several different perspectives starting from 2 up to 72 were fused at score-level for 4 different kinds of extracted features using a simple sum-rule based fusion. We achieved a best overall EER of 0.039% for the fusion of 18 different views and Maximum Curvature [34] features.

Yang et al. [35] proposed a decision-level fusion approach based on three different finger vein feature representations. They extracted a topological feature, a local moment based feature and a vein shape based feature. These features were compared individually by means of a nearest cosine classifier outputting the class which the input feature belongs to. These output decisions were then fused by the help of the Dempster-Shafer algorithm. The authors reported a lowest FAR of 0.0086% at a FRR of 1%.

### 10.5.1.2 Multi-modality Fusion Including Finger Veins

In addition to the single modality fusion approaches, several multi-modality fusion approaches including finger veins as one of the involved biometric traits were proposed. Table 10.2 gives an overview of these approaches, including the reference to the original publication, the fusion level, the involved biometric traits, the number of subjects in the dataset used, the reported performance (EER if not stated otherwise) and the year of publication. Most approaches fuse finger-related biometrics, including fingerprint, finger texture, finger shape, finger knuckle and finger veins. There are only two approaches involving other biometrics than finger-related ones. Razzak et al. [36] fused face and finger veins and He et al. [37] fused face, fingerprints and finger veins. Both applied score-level fusion. The number of involved traits varies between at least two and at most four. Fingerprint is the most prominent one [37–46] besides finger veins that is included in the fusion followed by finger texture [38, 43, 45, 47–49] as the second most prominent one and finger shape [42, 43, 50–52] as the third one. The majority of the approaches is based on feature-level and score-level fusion, there are only two decision-level fusion approaches compared to eight



**Table 10.2** Related work in finger vein fusion, multi-modality fusion involving finger veins, ordered according to fusion level and year of publication

References	Fusion level	Involved traits	Subjects	Performance (EER)	Year
[40]	Feature	Fingerprint, finger veins	40	1.85% FRR and 0.97% FAR	2011
[44]		Fingerprint, finger veins	64	1.35% FAR at 0% FRR	2012
[46]		Fingerprint, finger veins	40	1.485%	2012
[48]		Finger texture, finger veins	220	0.45%	2012
[49]		Finger texture, finger veins	220	0.435%	2014
[43]		Finger texture, finger shape, fingerprint, finger veins	100	0.00796%	2015
[45]		Finger texture, fingerprint, finger veins	300	0.415%	2016
[51]		Score	Finger shape, finger veins	816	0.075%
[37]	Face, fingerprint, finger veins		510	99.8% GAR at 0.01% FAR	2010
[36]	Face, finger veins		35	5% FAR and 92.4% GAR	2010
[47]	Finger texture, finger veins		312	0.08%	2012
[52]	Finger shape, finger veins		120	4%	2013
[50]	Finger shape, finger veins		492	1.78%	2014
[42]	Finger shape, fingerprint, finger knuckle, finger veins		100	0.0319%	2014
[38]	Finger texture, fingerprint, finger veins		378	0.109%	2015
[41]	Decision	Fingerprint, finger veins	33	1.86%	2011
[39]	Feature/decision	Fingerprint, finger knuckle, finger veins	165	0.04%	2016

feature-level and eight score-level ones. All proposed fusion approaches showed a significant improvement in the recognition accuracy of the fusion compared to using finger veins only.

## 10.6 Experimental Analysis

This section describes the experimental part of this chapter. At first, the used subset of the dataset introduced in Sect. 10.4 is explained. Afterwards, the finger vein recognition tool chain which is employed during the experimental analysis is described. This is followed by a presentation of the fusion strategy and the applied score-level fusion framework. Afterwards, the experimental protocol to determine the FAR and FRR and consequently the recognition performance in terms of EER/FMR1000/ZeroFMR is explained. Then the results of the individual fusion strategies are given and discussed. Finally, this section is concluded with an overall results discussion.

### 10.6.1 Finger Vein Dataset

To reduce the amount of data during the fusion, we used a subset of the multi-perspective finger vein dataset [2] only. Not all 360 different perspectives are evaluated, but only each fifth one is considered. Thus, there is a total of 73 different perspectives ( $\frac{360^\circ}{5^\circ/\text{step}} = 72$  plus the last one which is  $360^\circ = 0^\circ$  again results in 73). All 63 capture subjects, 4 fingers per subject and 5 images per view and finger are considered. This results in a total of  $73 \times 63 \times 4 \times 5 = 91,980$  images instead of 454,860 for the total dataset.

### 10.6.2 Finger Vein Recognition Tool chain

The finger vein recognition tool chain includes all steps of a biometric recognition system starting with the extraction of the Region of Interest (ROI) to preprocessing, feature extraction and comparison. The input data are the images of the different individual perspectives acquired from the 3D capture device, the output is a comparison score that can be used to determine whether the provided finger belongs to a certain (enrolled) data subject or not.

#### ROI Extraction

Prior to the ROI extraction, the finger is aligned and normalised. The alignment should place the finger always in the same position in the image, independent of the relative position of the finger during the acquisition. To achieve this, the finger lines (edge between finger and the background of the image) are detected and the centre

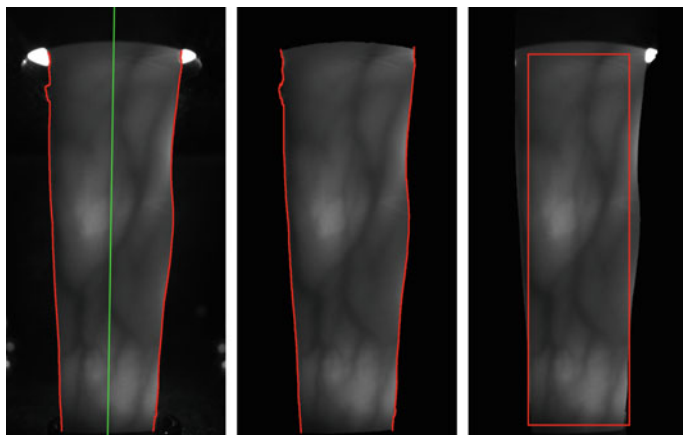


Fig. 10.7 ROI extraction process (images originally published in [2], ©2018 IEEE)

line (in the middle of the two finger lines) is determined. Afterwards, the centre line of the finger is rotated and translated in a way that it is placed in the middle of the image and the image region outside of the finger is masked by setting the pixels to black. The final step is to extract a rectangular ROI of a fixed size ( $1100 \times 300$  pixel) from a fixed position. The three steps are visualised in Fig. 10.7. The implementation used is based on the method proposed in [53].

### Preprocessing

Preprocessing tries to enhance the low contrast and improve the image quality. In the following the preprocessing methods, we employed in our finger vein recognition tool chain are explained.

Simple **CLAHE** [54] or other local histogram equalisation techniques are most prevalent according to the literature for this purpose. A localised contrast enhancement technique like CLAHE is a suitable baseline tool to enhance the vein images as they exhibit unevenly distributed contrast. CLAHE has an integrated contrast limitation (clip limit) which should avoid the amplification of noise.

**High-Frequency Emphasis Filtering (HFEF)** [55], originally proposed for hand vein image enhancement tries to enhance the vein images in the frequency domain. At first, the discrete Fourier transform of the image is computed, followed by the application of a Butterworth high-pass filter of order  $n$ . The authors originally proposed to use a global histogram equalisation but we decided to apply CLAHE instead.

**Circular Gabor Filter (CGF)** as proposed by Zhang and Yang [56] is another finger vein image enhancement technique which is rotation invariant and achieves an optimal joint localisation in both, the spatial and the frequency domain. The authors originally suggested to use grey level grouping for contrast enhancement but we again apply CLAHE instead.

Furthermore, the images were resized to half of their original size, which not only speeded up the comparison process but also improved the results. For more details on

the preprocessing methods, the interested reader is referred to the authors' original publications.

### Feature Extraction

We used five different feature extraction methods. The first three techniques discussed aim to extract the vein pattern from the background resulting in a binary image (vein pattern based methods) followed by a comparison of these binary images using a correlation measure. All algorithms are well-established finger vein recognition algorithms. We used the publicly available implementations published in [5].

**Maximum Curvature (MC [34])** aims to emphasise only the centre lines of the veins and is therefore insensitive to varying vein widths. The first step is the extraction of the centre positions of the veins by determining the local maximum curvature in cross-sectional profiles obtained in four directions: horizontal, vertical and the two oblique directions. The cross-sectional profile is determined based on the first and second derivatives. Then each profile is classified as either being concave or convex, where only the local maxima belonging to a concave profile indicate a vein line. Afterwards, a score according to the width and curvature of the vein region is assigned to each centre position and recorded in a matrix called locus space. Due to noise or other distortions, some pixels may not have been classified correctly at the first step, thus the centre positions of the veins are connected using a filtering operation in all four directions taking the 8-neighbourhood of pixels into account. The final binary output image is obtained by thresholding of the locus space using the median as a threshold.

**Principal Curvature (PC [57])**: At first the gradient field of the image is calculated. In order to prevent the unwanted amplification of small noise components, a hard thresholding which filters out small gradients by setting their values to zero is done. Then the gradient at each pixel is normalised to a magnitude of 1 to get a normalised gradient field. This normalised gradient field is smoothed by applying a Gaussian filter. The next step is the actual principal curvature calculation. The curvatures are obtained from the Eigenvalues of the Hessian matrix at each pixel. The two Eigenvectors of the Hessian matrix represent the directions of the maximum and minimum curvature and the corresponding Eigenvalues are the principal curvatures. Only the bigger Eigenvalue which corresponds to the maximum curvature among all directions is used. The last step is a threshold based binarisation of the principal curvature values to arrive at the binary vein output image.

**Gabor Filter (GF [47])**: Gabor filters are inspired by the human visual system's multichannel processing of visual information and have been widely used in biometrics. A Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave. Kumar and Zhou [47] proposed a Gabor filter based finger vein extraction approach. Therefore, a filter bank consisting of several 2D even symmetric Gabor filters with different orientations (in  $\frac{\pi}{k}$  steps where  $k$  is the number of orientations) is created.  $k$  feature images are extracted by filtering the vein image using the different filter kernels contained in the Gabor filter bank. The final feature image is obtained by summing all the single feature images from the previous step and thresholding

the resulting feature image. This image is then post-processed using morphological operations to remove noise to get the final binary vein output image.

In contrast to the vein pattern based techniques described above, two key-point based techniques were used. Key-point based techniques try to use information from the most discriminative points as well as considering the neighbourhood and context information around these points by extracting key-point locations and assigning a descriptor to each detected key-point location.

The first one is a **Scale-Invariant Feature Transform** (SIFT [58]) based technique with additional key-point filtering along the finger boundaries to suppress information originating from the finger shape instead of the vascular pattern. This technique was originally proposed by Kauba et al. [33].

**Deformation-Tolerant Feature Point Matching** (DTFPM [13]): The second key-point based technique replaces the conventional SIFT descriptor and key-point detector by vascular pattern tailored ones. This method is robust against irregular shading and vein deformations due to posture changes. At first, the authors apply a technique originally proposed by Yang and Yang [59] for enhancing the vein images. Then a minimum-curvature map is calculated from the enhanced vein images based on Eigenvalue analysis. The feature point locations are determined from this curvature image (smaller Eigenvalue) at any point where the vein shape is non-linear. The feature descriptor takes the vein shape around the key-point location into account and is extracted from the so-called vein pattern map (larger Eigenvalue). The feature vector contains a quantification of the different vein directions inside a variable-sized window around the key-point location. The descriptor is normalised with the help of a finger shape model in a way that the descriptor area becomes smaller the closer the key-point location is to the finger boundaries. The authors claim that their proposed method is tolerant against several different types of finger posture changes, e.g. longitudinal finger rotation, translations and bending of the finger.

## Comparison

For the comparison of the binary feature images we extended the approach in [25] and [34]. As the input images are neither registered to each other nor aligned, the correlation between the input image and in x- and y-direction shifted versions of the reference image is calculated. The maximum of these correlation values is normalised and then used as the final comparison score.

The SIFT features are compared by finding their nearest neighbours/best correspondences and calculating a score based on the distances between the corresponding key-points.

DTFPM employs a deformation tolerant comparison strategy by using non-rigid registration. At first, the correspondences between the key-points in the two images for comparison are found. These correspondences are filtered using a local and global histogram technique based on the relative distances between the corresponding key-points. After this filtering step, the key-point coordinates of one of the involved feature vectors are transformed by applying a non-rigid transformation based on an outlier-robust thin-plate spline model as proposed in [60]. Afterwards, the correspondences between the adjusted key-points are determined again. These updated

correspondences are filtered by a comparison of the descriptor distances with fixed thresholds. The final comparison score is determined as the ratio of the matched points and the sum of the number of detected key-points in both images.

### 10.6.3 Score-Level Fusion Strategy and Toolkit

We applied three different fusion strategies. The first strategy involves the fusion of all possible combinations of pairs of distinct views (which are  $\binom{N}{k} = \binom{73}{2} = 2628$  combinations, 73 different views are considered) as well as all possible three tuples of distinct views (which are  $\binom{73}{3} = 62196$  combinations) for each of the five-feature extraction methods. As motivated in the introduction, it is beneficial if the number of involved views is as little as possible to reduce the complexity and the production costs of the biometric capture device and to be able to build such a device without any moving parts. Thus, only pairs and three tuples are considered here. The second strategy employs the fusion of all possible combinations of feature extraction methods per view. There are  $\binom{5}{2} + \binom{5}{3} + \binom{5}{4} + \binom{5}{5} = 26$  combinations per perspective, resulting in a total of 10,830 different fusion combinations. Here, our aim is to identify the best combination of features for each individual view which does not necessarily have to be the same across all the different views. The third strategy is a combination (fusion) of the best results obtained during the first and second one.

All three fusion strategies are applied at score-level. The second strategy could be applied at feature-level too, but not for all the involved feature extraction types as they are not compatible with each other. The feature-level fusion of MC, PC and GF is possible while the fusion of DTFPM and SIFT with any of the other feature extraction types is not possible. Feature-level fusion is not possible for the first strategy at all, as there is no meaningful way to combine the features of different perspectives, e.g. by merging the extracted vein lines or using majority voting as the visible vein lines differ for each view. Score-level fusion usually performs better than decision-level fusion, as there is more information available at the score level and there are more variants to fuse the individual scores. Hence, we decided to apply score-level fusion in all three fusion strategies.

In our previous work [2], a simple sum based fusion rule, without any weights for the input scores, was applied. In this work, a more advanced score-level fusion approach, namely the BOSARIS toolkit [14] is utilised. BOSARIS provides a MATLAB based framework for calibrating, fusing and evaluating scores from binary classifiers and has originally been developed for automatic speaker recognition. It can be applied to any biometric trait where two alternate classes are distinguished (genuine/impostor). The toolkit provides several functionalities, e.g. a normalised Bayes error rate plot, ROC and DET plots, including efficient algorithms to generate these plots for large score files, logistic regression solutions for the fusion of several subsystems, solutions for calibration (mapping scores to likelihood ratios), a logistic regression optimiser and an efficient binary score file format. During this work, we only harness the fusion capabilities of BOSARIS though. BOSARIS needs

a supervised training phase where combination weights are trained based on logistic regression in order to fuse multiple input systems into a single output one providing well-calibrated log-likelihood-ratios. This is achieved by employing a general purpose, unconstrained convex optimisation algorithm, which is used to train the logistic regression fusion and calibration methods. Hence, BOSARIS needs a training set of data to find the optimal combination of weights for the actual fusion in order to minimise the classification error and thus to maximise the recognition performance based on the fused output scores. BOSARIS has the option to set a target prior according to the costs of a miss and a false alarm for the training phase of the fusion. We set this target prior to 0.5 assuming that the costs of a miss and a false alarm are both weighted equally.

### ***10.6.4 Evaluation Protocol***

The experiments are split into four parts: in the first part, we analyse the recognition performance of all single perspectives. Every perspective is considered as a separate dataset. Here, we do not perform any cross-projection comparison. The images are processed as described in Sect. 10.6.2 and 73 projections all around the finger in  $5^\circ$  steps are extracted. The recognition performance is quantified in terms of the EER as well as the FMR1000 (the lowest FNMR for FMR = 0.1%) and the ZeroFMR (the lowest FNMR for FMR = 0%). The performance values are calculated for each single perspective. For the parameter optimisation, the data set is divided into two roughly equal-sized subsets. The division is based on the contained subjects, i.e. all fingers of the same person are in one subset. Each subset is used to determine the parameters which are then applied to the other subset. This ensures a 100% separation of the data used for determining the optimal parameters and the actual test set. The necessary comparison scores for the FAR/FRR calculation, which is the basis for the EER/FMR1000/ZeroFMR calculation, are determined according to the test protocol of the FVC2004 [61]: to compute the genuine scores, all possible genuine comparisons are done. Instead of computing all possible impostor scores only the first image of a finger is compared against the first image of all other fingers. The final results are evaluated based on the combined scores (genuine and impostor) of both test runs. The parameter optimisation is executed only for the palmar dataset. The same parameter settings are also applied for the experiments on the other perspectives. The resulting number of comparisons for both subsets are listed in Table 10.3. All performance-related result values are given in percentage terms, e.g. 0.04 means 0.04%.

In the second part of our experiments, we fuse different features originating from the same feature extraction method but extracted from different perspectives as described in Sect. 10.6.3. The third part of the experiments is dedicated to a multi-algorithm fusion. We fuse all possible combinations of the five employed feature extraction methods at score level based on the scores obtained during the first part of the experiments, resulting in 2-, 3-, 4- and 5-tuples. In the last part, we com-

**Table 10.3** Number of comparisons for each subset

Name	Subjects	Genuine	Impostor	Total
Subset 1	32	1280	8128	9408
Subset 2	31	1240	7626	8866
Total	63	2520	15,754	18,274

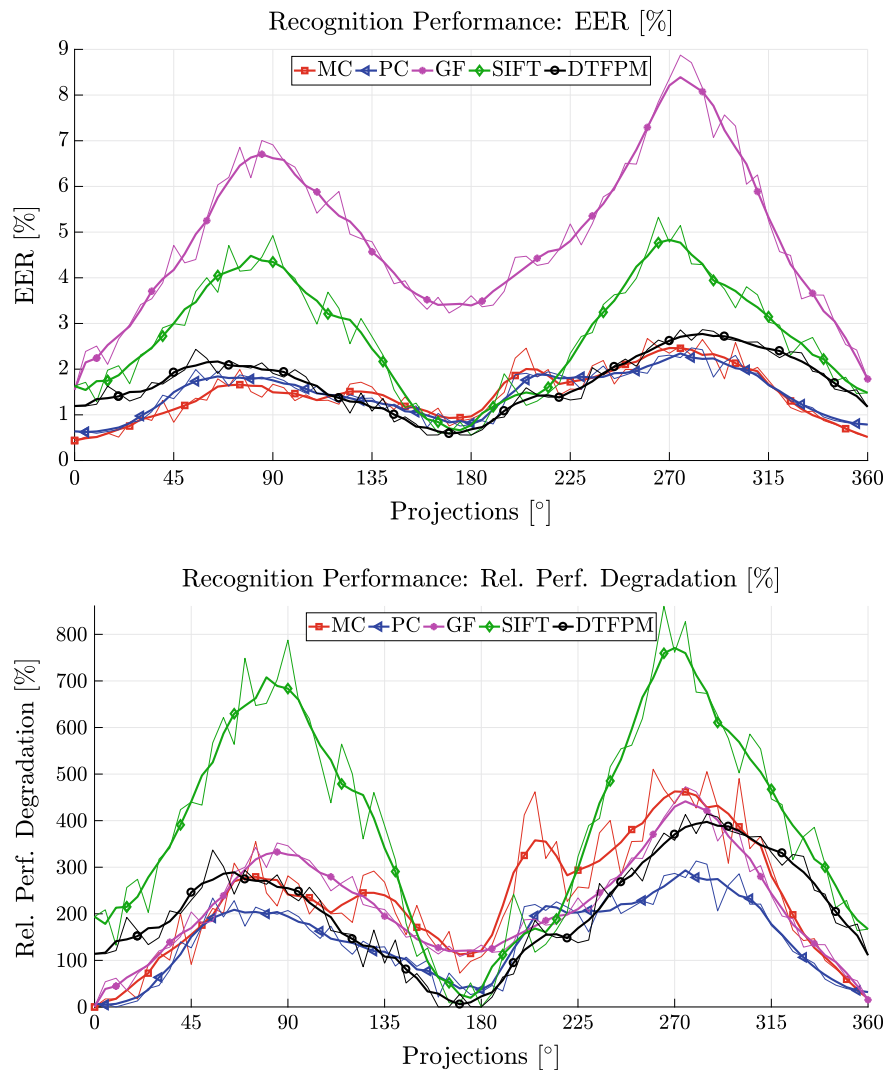
bine the two strategies of multi-perspective and multi-algorithm fusion. Based on the results from the two individual fusion strategies we determine the best possible combinations/fusion of perspectives and feature extraction methods. All four parts are evaluated using the same protocol to determine the performance figures. For all fusion experiments, the input data are the comparison scores generated during the single perspective experiments. We apply a fivefold cross-validations procedure, where we use every fold once for the training of the fusion module. The determined fusion parameters are applied to the test data consisting of the four remaining folds. The final results are evaluated based on the combined scores (genuine and impostor) of all five test runs.

We provide the scores files for each individual perspective and feature extraction methods as well as a script to run BOSARIS and generate all the fused scores files and performance figures we used during our experiments. These files and the scripts can be downloaded at <http://www.wavelab.at/sources/Prommegger19b/>.

### 10.6.5 *Single Perspective Performance Results*

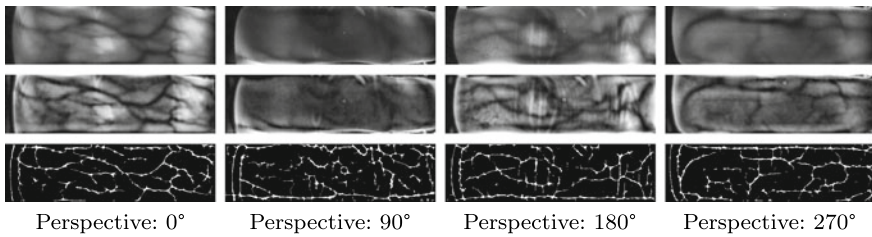
The single perspective analysis for MC, PC, GF and SIFT have already been carried out in our previous work [2]. We added DTFPM as an additional key-point based recognition scheme. We had to change our ROI extraction to make the ROIs compatible with DTFPM. Our previous ROI approach selected a fixed size rectangle placed at the centre of the finger, independent of the finger's width. DTFPM is sensitive to parts of the finger outline and background areas that are contained in the input images and expects the finger width normalised to the ROI height. Thus, we updated our ROI extraction scheme as described in Sect. 10.6.2 and recalculated the results for the already evaluated algorithms based on the new ROIs. Note that due to the new ROIs these updated results are different from our previous work. Figure 10.8 top shows the results in terms of EER. There are two lines for every method: the thin line shows the actual EER value, the thicker line is a smoothed version calculated based on the EER using a moving average filter of size 5, which should highlight the trend of the recognition performance. The images captured of neighbouring views contain quite a similar vein structures (note that our step-width is 5°), thus the recognition performance is similar too. The best results are obtained around the palmar (0°, 360°) and dorsal (180°) region. The results of the perspectives in-between are inferior. This





**Fig. 10.8** Recognition performance for different projections: EER (top) and relative performance degradation in relation to the best performing view (bottom)

is due to the fact, that they contain fewer visible vein lines and thus fewer vein information than the palmar and dorsal view. Figure 10.9 shows the original ROI, the ROI after preprocessing and the extracted features (using MC) for the views 0°, 90°, 180° and 270°. It reveals that the 90° and 270° views contain less vein information than the palmar and dorsal view. Moreover, the vein extraction algorithms include some features related with the texture of the finger. This is especially visible at 180° where



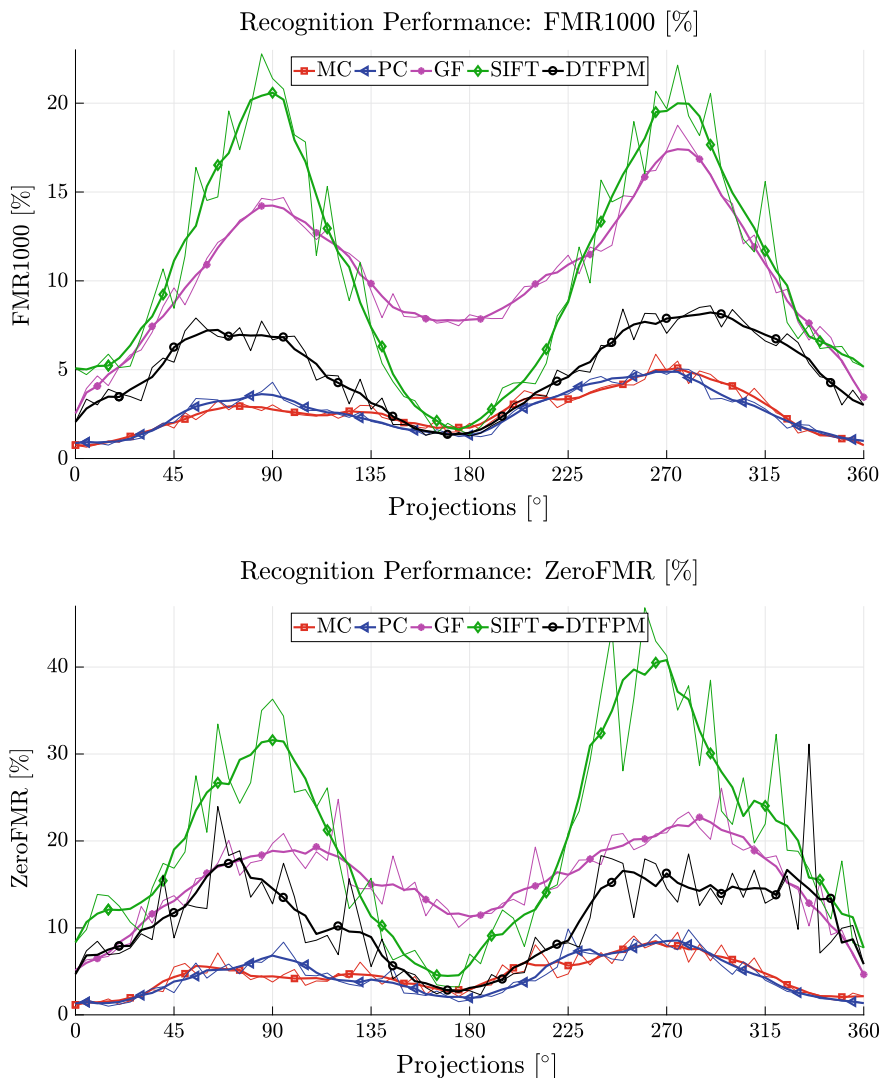
**Fig. 10.9** ROI (first row), enhanced images (second row) and extracted MC features (third row) for different projections (originally published in [2], ©2018 IEEE). Note that there are less vein lines visible for 90° and 270° compared to 0° and 180°

some of the features are related with the finger knuckles instead of veins. These features are visible as horizontal lines in the feature image.

For the key-point based algorithms, especially SIFT, the palmar region exhibits a better performance than the other perspectives as well, but the best performance is achieved around the dorsal region. For SIFT this can be explained based on the employed preprocessing: only image (vein) enhancement and no vein extraction (binarisation) ahead of the SIFT key-point calculation is applied. Hence, the non-vein finger texture information is not suppressed in the input images of SIFT. Especially, the structure of finger knuckles seem to contain a lot of additional information which SIFT is able to exploit during feature extraction. Finger knuckles have been introduced by Zhang et al. [62] as an independent biometric characteristic. Yang et al. [63] experienced a similar behaviour. They fused the finger texture of the dorsal view with the vein structure of the palmar view which leads to an improvement in the recognition performance. Consequently, the additional information originating from the finger knuckles and the finger texture present at the dorsal view leads to the superior performance of SIFT for the dorsal view compared to the palmar one.

Table 10.4 lists the information regarding the best and worst perspective for each feature extraction method. MC, PC and GF perform best around the palmar view (note that  $360^\circ = 0^\circ$ ), while SIFT and DTFPM perform best around the dorsal view. The overall best result was achieved for MC at  $0^\circ$  with an EER of 0.44% ( $\pm 0.15$ ) where the number in brackets is the confidence interval. For all feature extraction methods, the worst results can be reported around  $270^\circ$ . The Relative Performance Degradation (RPD) of the different perspectives is visualised in Fig. 10.8 bottom. The RPD, stated in Eq. (10.1), is calculated with respect to the minimum EER ( $EER_{\min}^{FT}$ ) reached for a certain feature extraction method, where  $EER_{perspective}^{FT}$  is the EER of the current perspective. The maximum performance degradation across the different algorithms is between 200 and 800%.

$$RPD_{perspective}^{FT} = \frac{EER_{perspective}^{FT} - EER_{\min}^{FT}}{EER_{\min}^{FT}} \quad (10.1)$$



**Fig. 10.10** Recognition performance among the different projections: FMR1000 (top), ZeroFMR (bottom)

The FMR1000 and ZeroFMR are visualised in Fig. 10.10 top and bottom, respectively. They follow the same trend as the EER: a good performance around the palmar and dorsal region and an inferior one for the views in between.

**Table 10.4** Best/worst single perspective results per feature extraction method and single perspective

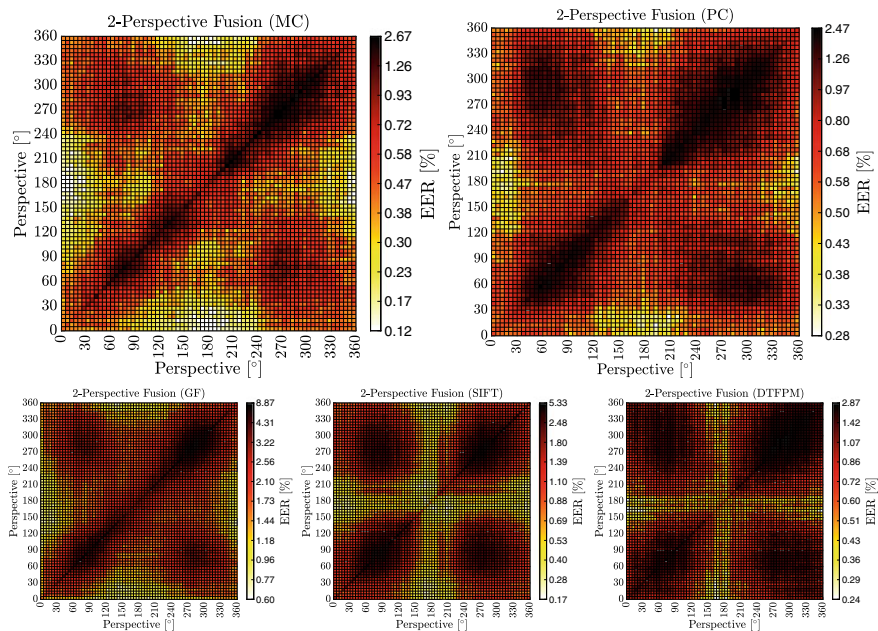
Feature type	Best perspective				Worst perspective			
	View	EER	FMR1000	ZeroFMR	View	EER	FMR1000	ZeroFMR
MC	0°	<b>0.44</b> (±0.15)	0.76	1.15	260°	2.67 (±0.37)	4.46	7.69
PC	10°	0.60 (±0.18)	0.87	1.35	280°	2.47 (±0.36)	5.02	9.79
GF	0°	1.55 (±0.28)	2.54	5.13	275°	8.87 (±0.65)	18.76	22.54
SIFT	180°	0.55 (±0.17)	1.35	6.98	265°	5.33 (±0.53)	20.67	42.98
DTFPM	160°	0.56 (±0.17)	1.31	3.13	285°	2.87 (±0.38)	8.51	12.56

### 10.6.6 Multi-perspective Fusion Results

In the second part of our experiments, we analyse the impact of fusing the extracted features of the same feature extraction method from multiple perspectives (MPF). In detail, we evaluate the fusion of all possible pairs and three tuples.

The first part of this section deals with the fusion of all possible pairs. Figure 10.11 shows heat maps of the EER for all combinations per feature extraction method (top row: MC, PC, bottom row: GF, SIFT and DTFPM). The perspectives involved in the fusion are plotted on x- and y-axis, whereas the performance in terms of EER is visualised using a colour scheme from light/white which corresponds to a low EER (good performance) to dark/red which corresponds to a high EER (bad performance). The actual logarithmic scale is given in the colour bar on the right side of the plots. Note that the results are symmetric with regard to the main diagonal (45°). This diagonal corresponds to the single perspective performance results and is visible as dark line (high EER) in all five plots.

According to the performance analysis of the single perspectives (Sect. 10.6.5), the palmar and dorsal region perform best. Although, there are slight variations among the different feature extraction methods, the results obtained from the single perspectives are confirmed by the two-perspective fusion: a combination of two perspectives including the palmar (close to 0°, 360°) or dorsal (close to 180°) region always results in a good recognition performance. A fusion of two views in-between those two regions result in an inferior performance. For MC, PC and GF the EER for all fusion combinations including the palmar (area along the outer edges of the plot) and dorsal view (cross lines in the centre) perform better (light, white to yellow colours) than fusion combinations without these views (dark, orange to red colours), achieving the best results when both regions are fused (light, white colour).



**Fig. 10.11** Recognition performance for two-view fusion. Top row: MC (left), PC (right), bottom row: GF (left), SIFT (middle) and DTFPM (right)

Both key-point based methods show a different behaviour. The fusion of the palmar and dorsal region is still superior to all other fusion combinations, but SIFT and DTFPM perform well if the dorsal perspective is included in the fusion in general. This can also be seen in the plots as the 180° cross shows light, white to yellow colours which indicates a good performance. For SIFT, this is even more pronounced than for DTFPM.

Table 10.5 lists the best results in terms of EER, FMR1000 and ZeroFMR for each feature extraction method in detail. MC when fusing 0° and 180° achieves the overall best performance with an EER of 0.12%. For the evaluation of the results, the single perspective baseline EER and the relative performance increase (RPI) with respect to the baseline EER, as calculated in Eq. (10.2), are stated. The performance increase compared to the best single view result is between 110% (PC) and 270% (MC), which corresponds to a 2–3.5 times lower EER than the single perspective performance, respectively.

$$RPI = \frac{EER_{Baseline} - EER_{Fusion}}{EER_{Fusion}} \tag{10.2}$$

In addition to all pairs, all possible triples are evaluated. Table 10.6 shows the five best performing combinations per feature extraction method. Again the single perspective baseline EER and the relative performance increase is included. The highest

**Table 10.5** Best two-perspective fusion results per feature extraction method. Best result is highlighted **bold font**. For comparability also the single perspective baseline EER and the relative performance improvement (based on the single perspective performance) is included

Feature type	2 Perspective fusion					Single perspective		Rel. Perf. Incr. [%]
	View 1	View 2	EER	FMR1000	ZeroFMR	View	EER	
MC	0°	180°	<b>0.12</b> (±0.08)	0.12	0.16	0°	0.44	<b>264.90</b>
PC	10°	190°	0.28 (±0.12)	0.36	0.56	10°	0.60	113.14
GF	140°	360°	0.60 (±0.18)	0.80	1.56	0°	1.55	156.48
SIFT	165°	205°	0.17 (±0.09)	0.36	1.63	180°	0.55	229.72
DTFPM	0°	160°	0.24 (±0.11)	0.32	1.55	160°	0.56	132.27

recognition performance improvement is between 150% for PC and 1100% for MC which is in any case better than the best two-perspective fusion (see Table 10.5). The overall best result with an EER of 0.036% is achieved using MC when fusing the 5°, 170° and 235° view.

Table 10.6 also includes the perspectives of interest. It is striking, that once again a lot of combinations include perspectives close to the palmar (0°, 360°) and dorsal (180°) regions. Thus, we additionally analysed the occurrence of the palmar and dorsal view in the top 25 results for each feature extraction method. All angles within a certain range around 0° and 180° are mapped to the palmar and dorsal region, respectively. Three different mapping ranges are evaluated: ±15° (345°–15°, 165°–195°), ±20° (340°–20°, 160°–200°) and ±25° (335°–25°, 155°–205°). The results are presented in Table 10.7. It turns out that the best performing individual region (palmar for MC, PC, GF and dorsal for SIFT and DTFPM) is present in most of the top 25 fusion combinations. At a mapping range of ±25° it is even included in at least 96% of the top 25 results. For this mapping range also the opposite region is part of at least 80% of the combinations, except for GF (only 24%). For GF, this can be explained by the big performance difference of palmar (~1.5%) and dorsal region (~3.6%).

In order to be able to decide whether a three-perspective fusion is beneficial compared to a two-perspective approach, one way is to calculate the significance of the recognition performance improvement. We use the method proposed in [64] to calculate a boundary for the significance from the achieved EERs. Table 10.8 lists the  $\chi^2$  values in detail. The following translations of  $\chi^2$  values into  $p_v$  values can be used to interpret the values stated in the table:  $\chi^2 = 6.6$  corresponds to  $p_v = 0.01$  ( $\equiv 1\%$ ),  $\chi^2 = 7.9$  to  $p_v = 0.005$  ( $\equiv 0.5\%$ ) and  $\chi^2 = 10.8$  to  $p_v = 0.001$  ( $\equiv 0.1\%$ ). Thus, all performance improvements exhibiting  $\chi^2 > 6.6$  are regarded as significant. The resulting  $\chi^2$  values indicate that a fusion of two and three perspectives lead to

**Table 10.6** Recognition performance for three-view fusion: five best results per feature extraction method. Best result per feature extraction method is highlighted **bold font**. For comparability also the single perspective baseline EER and the relative performance improvement (based on the single perspective performance) is included

Feature type	3 Perspective fusion						Single perspective		Rel. Perf.
	View 1	View 2	View 3	EER	FMR1000	ZeroFMR	View	EER	Impr. [%]
MC	5°	170°	235°	<b>0.036</b> (±0.04)	0.040	0.240	0°	0.44	<b>1111.78</b>
	0°	210°	235°	0.036 (±0.04)	0.040	0.120			1107.27
	10°	165°	215°	0.039 (±0.05)	0.040	0.159			1019.25
	20°	160°	235°	0.039 (±0.05)	0.040	0.040			1014.94
	165°	235°	355°	0.039 (±0.05)	0.040	0.159			1014.94
PC	10°	175°	200°	<b>0.238</b> (±0.11)	0.401	0.602	10°	0.60	<b>150.21</b>
	20°	205°	235°	0.239 (±0.11)	0.319	0.638			149.65
	175°	235°	360°	0.239 (±0.11)	0.399	0.518			149.65
	140°	190°	360°	0.239 (±0.11)	0.282	0.524			149.59
	155°	210°	360°	0.239 (±0.11)	0.399	0.839			149.45
GF	125°	225°	360°	<b>0.284</b> (±0.12)	0.401	1.325	0°	1.55	<b>446.48</b>
	90°	205°	360°	0.313 (±0.13)	0.638	1.794			394.98
	75°	140°	360°	0.321 (±0.13)	0.442	1.165			383.32
	120°	220°	355°	0.321 (±0.13)	0.758	1.475			383.09
	120°	200°	360°	0.321 (±0.13)	0.481	0.882			382.82
SIFT	165°	205°	350°	<b>0.058</b> (±0.05)	0.040	0.635	180°	0.55	<b>857.58</b>
	20°	170°	210°	0.075 (±0.06)	0.040	0.714			643.62
	170°	205°	350°	0.081 (±0.06)	0.079	0.476			585.30
	170°	205°	335°	0.081 (±0.06)	0.079	0.635			585.30
	140°	205°	350°	0.081 (±0.06)	0.079	0.714			585.30

(continued)

**Table 10.6** (continued)

Feature type	3 Perspective fusion						Single perspective		Rel. Perf.
	View 1	View 2	View 3	EER	FMR1000	ZeroFMR	View	EER	Impr. [%]
DTFPM	5°	160°	280°	<b>0.159</b> (±0.09)	0.559	1.837	160°	0.56	<b>249.88</b>
	0°	180°	295°	0.162 (±0.09)	0.439	1.276			243.31
	15°	160°	295°	0.162 (±0.09)	0.439	1.637			243.04
	0°	180°	185°	0.165 (±0.09)	0.437	1.033			237.24
	0°	180°	245°	0.169 (±0.09)	0.439	2.396			228.78

**Table 10.7** Analysis of the occurrence of palmar and dorsal views per feature extraction method in the 25 best three-perspective fusions. Both means that palmar and dorsal are present at the same combination.

Feature type (%)	Max distance ±15°			Max distance ±20°			Max distance ±25°		
	Palmar	Dorsal	Both	Palmar	Dorsal	Both	Palmar	Dorsal	Both
MC	84.0	52.0	40.0	92.0	76.0	68.0	<b>100.0</b>	84.0	84.0
PC	92.0	68.0	68.0	<b>100.0</b>	68.0	68.0	<b>100.0</b>	80.0	80.0
GF	<b>100.0</b>	8.0	8.0	<b>100.0</b>	16.0	16.0	<b>100.0</b>	24.0	24.0
SIFT	80.0	88.0	68.0	84.0	88.0	72.0	92.0	96.0	88.0
DTFPM	92.0	60.0	56.0	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>

a significant improvement compared to the single view performance, whereas the improvement for a three perspective fusion compared to fusing two views is lower but still significant for MC, GF and SIFT.

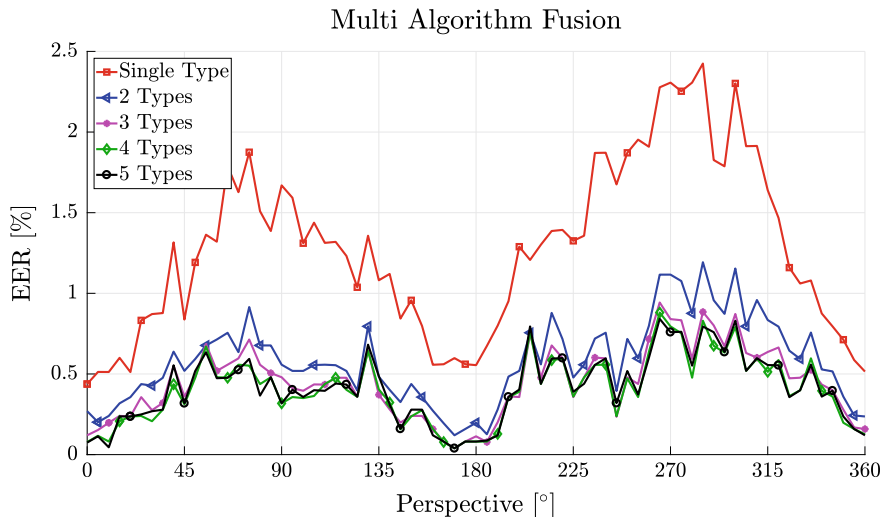
### 10.6.7 Multi-algorithm Fusion Results

This time different feature extraction methods per perspective are fused (MAF) instead of perspectives per feature extraction method. We evaluate all possible pairs, triples, quadruples and the combination of all five- feature extraction methods, resulting in 26 different combinations per perspective. Figure 10.12 shows the best fusion result per number of fused feature extraction methods. The best result, for example, two-feature extraction methods included in the fusion at 0° means that the best performing pair of features in terms of EER of all pairs calculated at 0° is depicted. It



**Table 10.8** Estimated  $\chi^2$  from the EER for multi-perspective fusion. Best results per number of involved views is highlighted **bold font**

Feature extraction method	Best EER for [n] involved views			Significance $n_1 \rightarrow n_2$ ( $\chi^2$ value)		
	n = 1	n = 2	n = 3	n = 1 $\rightarrow$ n = 2	n = 1 $\rightarrow$ n = 3	n = 2 $\rightarrow$ n = 3
MC	<b>0.44</b> ( $\pm 0.15$ )	<b>0.12</b> ( $\pm 0.08$ )	<b>0.036</b> ( $\pm 0.04$ )	33.415	62.660	8.265
PC	0.60 ( $\pm 0.18$ )	0.28 ( $\pm 0.12$ )	0.238 ( $\pm 0.11$ )	21.264	28.576	0.622
GF	1.55 ( $\pm 0.28$ )	0.60 ( $\pm 0.18$ )	0.284 ( $\pm 0.12$ )	<b>76.708</b>	<b>159.698</b>	<b>20.642</b>
SIFT	0.55 ( $\pm 0.17$ )	0.17 ( $\pm 0.09$ )	0.058 ( $\pm 0.05$ )	36.650	72.755	10.054
DTFPM	0.56 ( $\pm 0.17$ )	0.24 ( $\pm 0.11$ )	0.159 ( $\pm 0.09$ )	23.391	140.869	3.005



**Fig. 10.12** Recognition performance for multi-algorithm fusion: best result in terms of EER per number of feature extraction methods fused is depicted for each perspective

can be seen that even the fusion of two-feature extraction methods increases the performance remarkably. Adding the third feature extraction method further improves the result, whereas fusing four- or five-feature extraction methods does not further improve the recognition performance significantly.

Table 10.9 lists the results of the MAF in more detail. The column occurrence states how often in terms of perspectives a feature extraction method combination performs superior to all other combinations of the same number of included feature

**Table 10.9** Multi-algorithm fusion results per number of included features. Occurrence indicates the numbers of perspectives for which the specified combination achieves the best score, the given EER values are calculated over all perspectives. The two view columns state at which view the best and worst performance has been achieved. The best result per number of included feature extraction methods is highlighted **bold face**

# Features included	Feature types	Occurrences	Best		Avg	Worst	
			EER	View		EER	View
1	MC	34 (46.58%)	<b>0.44</b> (±0.15)	0°	1.46	2.67 (±0.37)	260°
	PC	19 (26.03%)	0.60 (±0.18)	10°	1.47	2.47 (±0.36)	280°
	DTFPM	16 (21.92%)	0.56 (±0.17)	160°	1.71	2.87 (±0.38)	285°
	SIFT	4 (5.48%)	0.55 (±0.17)	180°	2.75	5.33 (±0.53)	265°
	GF	–	1.55 (±0.28)	0°	4.89	8.87 (±0.65)	275°
2	PC, DTFPM	31 (42.47%)	0.20 (±0.10)	180°	0.66	1.32 (±0.26)	205°
	MC, DTFPM	22 (30.14%)	0.13 (±0.08)	185°	0.68	1.47 (±0.28)	285°
	MC, SIFT	11 (15.07%)	<b>0.12</b> (±0.08)	170°	0.78	1.83 (±0.31)	265°
	SIFT, DTFPM	8 (10.96%)	0.16 (±0.09)	175°	1.04	2.08 (±0.33)	265°
	MC, PC	1 (1.37%)	0.32 (±0.13)	10°	0.95	1.95 (±0.32)	285°
	PC, SIFT	–	0.24 (±0.11)	180°	0.92	1.88 (±0.31)	265°
	GF, DTFPM	–	0.32 (±0.13)	180°	1.17	2.32 (±0.35)	265°
	GF, SIFT	–	0.40 (±0.14)	170°	1.63	3.56 (±0.43)	265°
	MC, GF	–	0.44 (±0.15)	0°	1.39	2.54 (±0.36)	300°
	PC, GF	–	0.51 (±0.16)	360°	1.28	2.32 (±0.35)	265°

(continued)

**Table 10.9** (continued)

# Features included	Feature types	Occurrences	Best		Avg	Worst	
			EER	View	EER	EER	View
3	MC, SIFT, DTFPM	33 (45.21%)	<b>0.04</b> (±0.05)	170°	0.50	0.99 (±0.23)	285°
	MC, PC, DTFPM	23 (31.51%)	0.12 (±0.08)	185°	0.52	1.23 (±0.25)	205°
	PC, SIFT, DTFPM	11 (15.07%)	0.12 (±0.08)	165°	0.53	0.96 (±0.22)	270°
	PC, GF, DTFPM	3 (4.11%)	0.23 (±0.11)	245°	0.62	1.31 (±0.26)	205°
	MC, GF, DTFPM	2 (2.74%)	0.16 (±0.09)	185°	0.66	1.47 (±0.28)	285°
	MC, PC, SIFT	1 (1.37%)	0.12 (±0.08)	170°	0.64	1.31 (±0.26)	265°
	MC, GF, SIFT	–	0.12 (±0.08)	170°	0.77	1.76 (±0.30)	265°
	GF, SIFT, DTFPM	–	0.12 (±0.08)	175°	0.82	1.68 (±0.30)	265°
	PC, GF, SIFT	–	0.25 (±0.11)	170°	0.82	1.71 (±0.30)	265°
	MC, PC, GF	–	0.32 (±0.13)	0°	0.94	1.91 (±0.31)	285°
4	MC, PC, SIFT, DTFPM	51 (69.86%)	<b>0.04</b> (±0.05)	170°	0.42	0.88 (±0.21)	265°
	MC, PC, GF, DTFPM	10 (13.70%)	0.12 (±0.08)	185°	0.51	1.23 (±0.25)	205°
	MC, GF, SIFT, DTFPM	9 (12.33%)	0.04 (±0.05)	170°	0.50	1.07 (±0.24)	275°
	PC, GF, SIFT, DTFPM	3 (4.11%)	0.11 (±0.08)	185°	0.50	1.00 (±0.23)	265°
	MC, PC, GF, SIFT	–	0.09 (±0.07)	170°	0.63	1.32 (±0.26)	265°
5	MC, PC, GF, SIFT, DTFPM	73 (100.00%)	<b>0.04</b> (±0.05)	170°	0.41	0.84 (±0.21)	265°

extraction methods. The minimum, average and maximum EER are determined based on the results for all perspectives of the given feature extraction method combination. Considering single feature extraction methods, MC or PC are included in more than 70% of the best results. GF is not included in any combination that performs best for any perspective. The results of fusing feature extraction method pairs clearly show that it is beneficial to fuse a vein pattern based algorithm (MC, PC, GF) to a key-point based one (SIFT, DTFPM). The combinations of either MC/PC and SIFT/DTFPM are leading to 98% of the best results in two-feature extraction methods fusion.

**Table 10.10** Estimated  $\chi^2$  from the EER for multi-algorithm fusion

Nr of features EER	n = 1 0.44 ( $\pm 0.15$ )	n = 2 0.12 ( $\pm 0.08$ )	n = 3 0.04 ( $\pm 0.05$ )	n = 4 0.04 ( $\pm 0.05$ )	n = 5 0.04 ( $\pm 0.05$ )
n = 1 0.44 ( $\pm 0.15$ )	–	33.42	60.91	60.91	60.91
n = 2 0.12 ( $\pm 0.08$ )	33.42	–	7.31	7.31	7.31
n = 3 0.04 ( $\pm 0.05$ )	60.91	7.31	–	0	0
n = 4 0.04 ( $\pm 0.05$ )	60.91	7.31	0	–	0
n = 5 0.04 ( $\pm 0.05$ )	60.91	7.31	0	0	–

DTFPM (83%) is involved more often than SIFT (26%). Again, GF is not present in any of the best combinations. The overall best result with an EER of 0.04% is achieved when fusing MC, PC, SIFT and DTFPM. Once again, the analysis of the perspective, at which the best result is achieved, confirms, that views from the palmar ( $0^\circ$ ,  $360^\circ$ ) and dorsal ( $180^\circ$ ) region perform best.

Same as for the two-perspective fusion, we also check the performance increase of three-perspective fusion on its significance. Table 10.10 lists the results in detail. The resulting  $\chi^2$  values indicate, that a fusion of two or more feature extraction methods is always beneficial compared to a single feature extraction method. The same holds true when comparing a two-feature extraction method fusion to a three, four or five one. However, applying a four or five feature-type fusion instead of a three feature-type one leads to no significant improvements anymore.

### 10.6.8 Combined Multi-perspective and Multi-algorithm Fusion

In this section, we combine multiple perspectives and multiple feature extraction methods into one combined fusion method (CMPMAF). For the selection of the relevant perspectives and feature extraction methods we considered the results for multi-perspective fusion (Sect. 10.6.6) and feature extraction method fusion (Sect. 10.6.7). Although the  $\chi^2$  values for the multi-perspective fusion in Table 10.8 are only boundaries, they still indicate that the performance increase from two to three perspectives is significant for MC, GF and SIFT. The drawback of adding additional perspectives is the added cost/complexity to the system (additional camera and illumination module, higher computational costs). Therefore, we decided that the significance of the improvement is not high enough to justify the extra effort. As

**Table 10.11** Performance results: Fusion of vein pattern based with key-point based features for both, palmar and dorsal view. The best result is highlighted **bold face**

Feature types	Perspectives	EER	FMR1000	ZeroFMR
MC, SIFT	0°, 180°	<b>0.04 (±0.05)</b>	0.04	0.64
MC, DTFPM	0°, 180°	0.08 (±0.07)	0.08	0.12
PC, SIFT	0°, 180°	0.16 (±0.09)	0.16	0.32
PC, DTFPM	0°, 180°	0.16 (±0.09)	0.16	0.24
GF, SIFT	0°, 180°	0.20 (±0.10)	0.20	0.60
GF, DTFPM	0°, 180°	0.20 (±0.10)	0.20	0.28

a result of this, we only consider the two perspective fusion. The results presented in Fig. 10.11 and Table 10.5 show that the best results are achieved when fusing palmar and dorsal view. This behaviour can be confirmed when analysing the occurrence of certain perspectives of the three-perspective fusion: Table 10.7 states that the palmar and dorsal region is part of most of the top 25 results. Therefore, we selected 0° and 180° for our combined fusion.

For MAF, the significance analysis (see Table 10.10) indicates that the performance increase from a two to a three feature extraction method fusion is significant but would lead to additional computational costs (for score-level fusion, every feature extraction method needs to be processed by the whole processing chain up to the comparison). Thus, we decided to include the two-feature extraction method MAF into our combined fusion strategy only. Furthermore, the results listed in 10.9 state that 88% of the best two-feature extraction method fusion combinations include one vein pattern based (MC, PC, GF) and one key-point based (SIFT, DTFPM) feature. Therefore, we analysed all possible combinations of those feature extraction methods using both, palmar and dorsal view. Table 10.11 lists the results of the CMPMAF. We evaluated all six possible combinations and arrived at a best EER of 0.04% with a confidence interval of 0.05% for the combined fusion of MC and SIFT for palmar and dorsal view. This result is 11 times better than the best single perspective result (MC at 0° with an EER of 0.44%). All other combinations also perform well. The worst result with an EER of 0.20% is achieved when fusing GF with either SIFT or DTFPM. This is still more than two times better than the best single perspective result. For the sake of completeness, we also calculated the results of the best 3-, 4- and 5-MAF combinations with the palmar and dorsal view. These results, listed in Table 10.12, show that the EER can be further improved. The best result with an EER of 0 is achieved when fusing the scores of all five feature types.

Table 10.13 compares the performance of the best combined two-perspective two-algorithm fusion with the best results of all other fusion strategies. One can see that the calculated  $\chi^2$  indicates a significant performance improvement with respect to the single perspective, the 2-MPF and the 2-MAF strategy. All other fusion strategies achieved about the same EER.

**Table 10.12** Performance results: Fusion of vein pattern based with key-point based features for both, palmar and dorsal view. The best result is highlighted **bold face**

Feature types	Perspectives	EER	FMR1000	ZeroFMR
MC, SIFT, DTFPM	000°, 180°	0.04 ( $\pm 0.04$ )	0.00	0.36
MC, PC, SIFT, DTFPM	000°, 180°	0.01 ( $\pm 0.01$ )	0.00	0.12
MC, PC, GF, SIFT, DTFPM	000°, 180°	<b>0.00 (<math>\pm 0.00</math>)</b>	0.00	0.00

**Table 10.13** Comparison of the best two-perspective two-algorithm fusion combination to the best result of the other fusion strategies including the relative performance improvement, the factor, by which the EER decreased and the boundary  $\chi^2$  for significance

Fusion strategy	EER	EER CMPMAF	Rel. Perf. Impr. [%]	Factor	$\chi^2$
Single perspective	0.44 ( $\pm 0.15$ )	0.04 ( $\pm 0.05$ )	1000	11	60.91
2-MPF	0.12 ( $\pm 0.08$ )		200	3	7.31
3-MPF	0.04 ( $\pm 0.04$ )		0	1	0.00
2-MAF	0.12 ( $\pm 0.08$ )		200	3	7.31
3-MAF	0.04 ( $\pm 0.05$ )		0	1	0.00
4-MAF	0.04 ( $\pm 0.05$ )		0	1	0.00
5-MAF	0.04 ( $\pm 0.05$ )		0	1	0.00

### 10.6.9 Results Discussion

The evaluation of the independent recognition performances for different projections revealed, that indeed the widely used palmar perspective performed best, followed by the dorsal one performing second best. The views in-between exhibit a slightly worse performance, which is still acceptable. Our results indicate that the presence of finger texture and finger knuckles has a positive influence on the recognition performance. Figure 10.9 shows, that the well-established feature extraction algorithms not only extract features resulting from the finger veins but also from the skin texture of the finger and therefore inherently fuse texture and vein structure. The best single view result was achieved using MC features at the palmar view with an EER of 0.44%.

However, the main objective of this work was to find a suitable trade-off between the number of involved views and feature extraction methods and the recognition performance. In order to arrive at a design decision for a multi-perspective finger vein capture device, several aspects have to be considered: first of all, the gain in recognition accuracy, followed by the production costs and complexity of the biometric capture device which is directly related to the number of involved views and finally the computational complexity of the finger vein recognition system including

the capturing time, i.e. the total processing time, which is related to both, the number of different views and the number of different feature extraction methods involved. Adding more perspectives or feature extraction methods increases the complexity of the finger vein sensor and the recognition tool chain. For every feature extraction method, all steps of the recognition tool chain from preprocessing to comparison need to be executed. Adding further perspectives additionally increases the cost and complexity of the capture device's hardware by the need of either adding more camera/illumination modules (one per perspective) or a rotator that moves camera and illumination module into position. Ideally, the number of perspectives and feature extraction methods are kept to a minimum. Furthermore, additional aspects like an improved resistance against presentation attacks and an increased robustness against environmental influences should be included too. Therefore, the decision on how many perspectives and feature extraction methods are used has to be a trade-off between added cost/complexity and improvement of the recognition performance. Our proposed design is based on the findings during the fusion evaluations.

The multi-perspective fusion results showed that by fusing two independent views, in particular, the palmar and dorsal view, a significant performance gain can be achieved. Adding a second perspective improved the recognition performance between a factor 2–3.5, depending on the feature extraction method. The best result with an EER of 0.12% was achieved using MC features fusing the palmar and dorsal view. Adding a third view still improves the performance compared to two perspectives, but not to the same extent (significance) as from a single perspective to the 2-MPF. In this case, the best result of 0.036% EER was achieved using MC when fusing 5°, 170° and 235°. A biometric capture device able to capture the palmar and the dorsal view simultaneously can be built without any moving parts. Two cameras and two illumination modules are sufficient. Each additional view poses noticeable extra costs in terms of hardware (camera and illumination modules) and complexity of the capture device construction. Therefore, one must decide whether the improvement in accuracy justifies the extra effort. As our results show, the performance improvement from a 2-MPF to a 3-MPF is not as significant as from a single perspective to a 2-MPF, a two-perspective capture device, capturing the vein structure from the palmar and dorsal region is the best choice.

For MAF, a single perspective capturing device is sufficient. Such a biometric capture device can be built in a more compact and less expensive manner than a multi-perspective one. Moreover, existing finger vein capture devices acquiring images of the palmar view, can be utilised to apply multi-algorithm fusion too. However, adding an additional feature type to the MAF increases the computational cost. The MAF results showed, that the fusion of different feature extraction methods per single view improves the overall performance remarkably as well. The best results were obtained when fusing vein pattern based algorithms (especially MC and PC) with key-point based methods (SIFT, DTFPM). The best MAF result with an EER of 0.04% was achieved when fusing MC, SIFT and DTFPM in the dorsal region. Including more feature types does not improve the performance compared to the 3-MAF. As the computational complexity for the calculation and comparison of DTFPM features are higher than for the other features types, and the performance increase compared

to the best 2-MAF utilising MC and SIFT (EER = 0.12%) features is not as significant as from a single perspective to the 2-MAF, the best MAF option is a 2-MAF including MC and SIFT features.

In a third step, we combined MPF and MAF. By using the best performing perspectives of the two-perspective approach (palmar and dorsal) and combining them with a vein pattern based (MC, PC or GF) and a key-point based method (SIFT or DTfPM), we were able to achieve an EER of 0.04% utilising MC and SIFT. This corresponds to an improvement by a factor of 11 compared to the best single perspective performance, while achieving similar results as for the best MPF and MAF strategies. Adding more feature types to the combined fusion strategy further improved the result. Combining palmar and dorsal view together with all five feature types resulted in a perfect result with EER, FMR1000 and ZeroFMR of 0%.

A multi-perspective finger vein capture device is more resistant against presentation attacks, especially against simple paper printout based attacks. Depending on the actual construction of the multi-perspective capture device, it might also be more robust against contamination (e.g. dust and dirt, sun protection lotion or hand cream on the finger surface) of the finger due to the fact that more than one perspective is captured. Hence, the two-perspective capture device is the preferred option over the single perspective, multi-algorithm fusion one regarding these additional aspects.

Taking all the above-mentioned considerations into account, especially the additional advantages provided by a multi-perspective capture device in terms of resistance against presentation attack and robustness against external influences, the most preferable option is to design a two-perspective capture device capturing the palmar and the dorsal view applying a two-algorithm fusion including MC and SIFT features, whereas by including only one view the advantages of multi-perspective recognition can not be retained. The second feature extraction method can be included without involving additional hardware costs just by extending the recognition tool chain and putting up with the extended processing time, which makes the two-feature version beneficial in any case. This proposed finger vein capture device set-up arrives at an EER of 0.04%, which is a performance gain by a factor of 11 compared to the best single-view, single feature performance. Hence, this option provides an optimal trade-off between recognition accuracy, construction costs and processing time.

## 10.7 Conclusion and Future Work

In this chapter, we introduced multi-perspective finger vein recognition. For most work reported in the literature, only the palmar view is used in finger vein recognition. However, as the finger is an elliptically shaped cylinder, there are several other views available all around the finger's longitudinal axis. In order to be able to exploit these additional views, a suitable biometric capture device able to capture these different views is necessary. This chapter is based on our previous work [2], where we constructed a rotating, multi-perspective finger vein capture device which was then utilised to capture a multi-perspective finger vein data set. Based on this dataset, the



recognition performance of each view was evaluated individually. Then we applied three different score-level fusion strategies, the first one fusing all possible pairs and triples of distinct views, the second one fusing all different feature combinations per each single view and the third one combining the first two approaches. The first strategy was employed to find out the best performing pairs and three tuples of views in terms of recognition performance. The more views are desired to be captured, the higher the complexity and production costs of a suitable biometric capture device. At some point (a certain number of desired views), only a rotating device is able to capture the desired views. A rotating capture device bears several disadvantages, e.g. it is more prone to failures and has an increased capturing time. If only a limited number of views is involved, the production costs and the complexity of the biometric capture device are kept low. The second strategy was applied to investigate the best feature extraction method combination per view. The third strategy, which combines the first two approaches, was applied to find out if the recognition results can be further improved.

The single view evaluation results confirmed that the widely used palmar perspective, followed by the dorsal one (not taking views which are only a few degrees off from the palmar and dorsal view into account), achieves the best performance in finger vein recognition. All the perspectives in-between the palmar and dorsal one exhibit an inferior recognition performance to the palmar and dorsal one. Regarding the multi-perspective score-level fusion it turned out that a fusion of only two perspectives increases the recognition performance significantly, where a fusion of the palmar and the dorsal view performed best. Adding a third perspective still improves the results over the two perspective ones, but not to the same extent as the two perspective ones. The multi-algorithm fusion achieves similar results to the multi-perspective one, arriving at an EER of 0.04% for the combination of three-feature extraction methods. A pure multi-algorithm fusion is preferable in terms of hardware costs and capture device's complexity but does not exhibit the advantages of a multi-perspective recognition in regards to resistance against presentation attacks and increased robustness against external influences. By applying both fusion approaches at the same time for the best performing two perspectives (palmar and dorsal) and the best performing two distinct feature extraction methods (MC, a vein pattern based one and SIFT, a key-point based one), we were able to improve the recognition performance by a factor of 11 compared to the best single view result, achieving an EER of 0.04%.

Regarding recognition performance, hardware costs, processing time and robustness against presentation attacks and external influences the overall best option is to go for the combined multi-perspective and multi-algorithm fusion. In particular, a finger vein capture device capturing the palmar and the dorsal view including MC and SIFT features in a combined fusion provides the best trade-off between the above mentioned considerations and is, therefore, our preferred design decision.

### Future Work

The first step will be the construction of a combined multi-perspective and multi-algorithm type fusion finger vein capture device to prove its applicability in real-life

applications of finger vein recognition. We plan to do extended tests with this device, regarding presentation attacks, robustness against external influences like changing ambient conditions as well as subject-related influences.

Besides the capture device construction, our future work will include further analysis using our multi-perspective finger vein dataset. There are several other aspects besides the single perspective performance and the fusion of multiple perspectives which can be evaluated based on this dataset. One example is the robustness evaluation of different finger vein recognition algorithms against longitudinal finger rotation, which we already performed in a separate work [65]. We showed that this kind of rotation poses a severe problem for most algorithms. Since for our dataset the longitudinal rotation angle is known, we will test different techniques to compensate the finger rotation, either by estimating the rotation angle based on the captured images only or by using the known rotation angle and then applying a rotation compensating transform.

Another interesting question is if the best performing view is consistent across different subjects/fingers. To perform this analysis we will extend our dataset to contain at least 100+ subjects and then conduct a subject/finger based analysis to find out if the palmar perspective is the best one for all or at least a majority of the subjects/fingers or if there are significant differences.

Another field of interest is finger vein recognition in the 3D space. Therefore, we want to reconstruct a 3D model of the finger vein structure based on multiple images captured in different perspectives and apply different feature extraction and comparison strategies.

**Acknowledgements** This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700259, project PROTECT—Pervasive and UseR Focused BiomeTrics BordEr ProjeCT.

## References

1. Kono M (2000) A new method for the identification of individuals by using of vein pattern matching of a finger. In: Proceedings of fifth symposium on pattern measurement, Yamaguchi, Japan, pp 9–12
2. Prommegger B, Kauba C, Uhl A (2018) Multi-perspective finger-vein biometrics. In: Proceedings of the IEEE 9th international conference on biometrics: theory, applications, and systems (BTAS2018), Los Angeles, California, USA, pp 1–9
3. Raghavendra R, Busch C (2015) Exploring dorsal finger vein pattern for robust person recognition. In: 2015 international conference on biometrics (ICB), pp 341–348
4. University of Reading. PROTECT multimodal DB dataset, June 2017. <http://projectprotect.eu/dataset/>
5. Kauba C, Prommegger B, Uhl A (2018) Focussing the beam—a new laser illumination based data set providing insights to finger-vein recognition. In: Proceedings of the IEEE 9th international conference on biometrics: theory, applications, and systems (BTAS2018), Los Angeles, California, USA, pp 1–9

6. Kauba C, Prommegger B, Uhl A (2019) Openvein—an open-source modular multi-purpose finger-vein scanner design. In: Uhl A, Busch C, Marcel S, Veldhuis R (eds) *Handbook of vascular biometrics*. Springer Science+Business Media, Boston, MA, USA, pp 77–112
7. Gray H, Goss CM (1974) Anatomy of the human body. *Am J Phys Med Rehabil* 53(6):293
8. Tome P, Marcel S (2015) On the vulnerability of palm vein recognition to spoofing attacks. In: *The 8th IAPR international conference on biometrics (ICB)*
9. Tome P, Vanoni M, Marcel S (2014) On the vulnerability of finger vein recognition to spoofing
10. Lu Y, Yoon S, Park DS (2014) Finger vein identification system using two cameras. *Electron Lett* 50(22):1591–1593
11. Zhang Q, Zhou Y, Wang D, Hu X (2013) Personal authentication using hand vein and knuckle shape point cloud matching. In: *2013 IEEE Sixth international conference on biometrics: theory, applications and systems (BTAS)*. IEEE, pp 1–6
12. Qi Y, Zhou Y, Zhou C, Hu X, Hu X (2016) Vein point cloud registration algorithm for multi-pose hand vein authentication. In: *2016 IEEE international conference on identity, security and behavior analysis (ISBA)*. IEEE, pp 1–6
13. Matsuda Y, Miura N, Nagasaka A, Kiyomizu H, Miyatake T (2016) Finger-vein authentication based on deformation-tolerant feature-point matching. *Mach Vis Appl* 27(2):237–250
14. Brümmer N, De Villiers E (2013) The bosaris toolkit: theory, algorithms and code for surviving the new def.arXiv:1304.2865
15. Uhl A (2019) State-of-the-art in vascular biometrics. In: Uhl A, Busch C, Marcel S, Veldhuis R (eds). In: *Handbook of vascular biometrics*. Springer Science+Business Media, Boston, MA, USA, 3–62
16. Changzhou Songyang Machinery & Electronics New Technic Institute (2018) SY42STH47-1684A high torque hybrid stepping motor data sheet. <https://www.pololu.com/file/0J714/SY42STH38-1684A.pdf>. Accessed 20 June 2018
17. IDS Imaging Development Systems GmbH (2018) UI-ML1240-NIR NIR-enhanced industrial camera data sheet. [https://de.ids-imaging.com/IDS/datasheet\\_pdf.php?sku=AB00184](https://de.ids-imaging.com/IDS/datasheet_pdf.php?sku=AB00184). Accessed 20 June 2018
18. Fujifilm Corporation (2018) Fujifilm HF9HA-1B product page. [http://www.fujifilmusa.com/products/optical\\_devices/machine-vision/2-3-15/hf9ha-1b/index.html](http://www.fujifilmusa.com/products/optical_devices/machine-vision/2-3-15/hf9ha-1b/index.html). Accessed 20 June 2018
19. Midopt Corporation (2018) MIDOPT LP780 NIR pass-through filter product page. <http://midopt.com/filters/lp780/>. Accessed 20 June 2018
20. International Electrotechnical Commission et al (2015) Multimodal and other multibiometric fusion. ISO/IEC TR 24722
21. Yang J, Jia Y (2012) A method of multispectral finger-vein image fusion. In: *2012 IEEE 11th international conference on signal processing (ICSP)*, vol 1. IEEE, pp 753–756
22. Guan F, Wang K, Mo H, Ma H, Liu J (2009) Research of finger vein recognition based on fusion of wavelet moment and horizontal and vertical 2dpc. In: *2nd international congress on image and signal processing, 2009 CISP'09*. IEEE, pp 1–5
23. Yang J, Zhang X (2010) Feature-level fusion of global and local features for finger-vein recognition. In: *2010 IEEE 10th international conference on signal processing (ICSP)*. IEEE, pp 1702–1705
24. Gupta P, Gupta P (2015) An accurate finger vein based verification system. *Digital Signal Process* 38:43–52
25. Miura N, Nagasaka A, Miyatake T (2004) Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification. *Mach Vis Appl* 15(4):194–203
26. Kauba C, Piciuccio E, Maiorana E, Campisi P, Uhl A (2016) Advanced variants of feature level fusion for finger vein recognition. In: *Proceedings of the international conference of the biometrics special interest group (BIOSIG'16)*, Darmstadt, Germany, pp 1–12
27. Zhou Y, Kumar A (2010) Contactless palm vein identification using multiple representations. In: *2010 fourth IEEE international conference on biometrics: theory applications and systems (BTAS)*. IEEE, pp 1–6

28. Yang Y, Yang G, Wang S (2012) Finger vein recognition based on multi-instance. *Int J Digital Content Technol Appl* 6(11)
29. Park KR (2012) Finger vein recognition by combining global and local features based on svm. *Comput Inf* 30(2):295–309
30. Liu F, Yang G, Yin Y, Xi X (2013) Finger-vein recognition based on fusion of pixel level feature and super-pixel level feature. In: *Biometric recognition*. Springer, pp 274–281
31. Qin H, Qin L, Xue L, He X, Chengbo Y, Liang X (2013) Finger-vein verification based on multi-features fusion. *Sensors* 13(11):15048–15067
32. Lu Y, Yoon S, Park DS (2013) Finger vein recognition based on matching score-level fusion of gabor features. *J Korean Inst Commun Inf Sci* 38(2):174–182
33. Kauba C, Reissig J, Uhl A (2014) Pre-processing cascades and fusion in finger vein recognition. In: *Proceedings of the international conference of the biometrics special interest group (BIOSIG'14)*, Darmstadt, Germany, Sept 2014
34. Miura N, Nagasaka A, Miyatake T (2007) Extraction of finger-vein patterns using maximum curvature points in image profiles. *IEICE Trans Inf Syst* 90(8):1185–1194
35. Yang J, Shi Y, Yang J, Jiang L (2009) A novel finger-vein recognition method with feature combination. In: *2009 16th IEEE international conference on image processing (ICIP)*. IEEE, pp 2709–2712
36. Razzak MI, Yusof R, Khalid M (2010) Multimodal face and finger veins biometric authentication. *Sci Res Essays* 5(17):2529–2534
37. He M, Horng SJ, Fan P, Run RS, Chen RJ, Lai JL, Khan MK, Sentosa KO (2010) Performance evaluation of score level fusion in multimodal biometric systems. *Pattern Recogn* 43(5):1789–1800
38. Kang W, Chen X, Qiuxia W (2015) The biometric recognition on contactless multi-spectrum finger images. *Infrared Phys Technol* 68:19–27
39. Khellat-Kihel S, Abrishambaf R, Monteiro JL, Benyettou M (2016) Multimodal fusion of the finger vein, fingerprint and the finger-knuckle-print using kernel fisher analysis. *Appl Soft Comput* 42:439–447
40. Lin K, Han F, Yang Y, Zhang Z (2011) Feature level fusion of fingerprint and finger vein biometrics. In: *International conference in swarm intelligence*. Springer, pp 348–355
41. Park YH, Tien DN, Lee HC, Park KR, Lee EC, Kim SM, Kim HC (2011) A multimodal biometric recognition of touched fingerprint and finger-vein. In: *2011 international conference on multimedia and signal processing (CMSP)*, vol 1. IEEE, pp 247–250
42. Peng J, El-Latif AAA, Li Q, Niu X (2014) Multimodal biometric authentication based on score level fusion of finger biometrics. *Optik-Int J Light Electron Opt* 125(23):6891–6897
43. Peng J, Li Q, El-Latif AAA, Niu X (2015) Linear discriminant multi-set canonical correlations analysis (ldmcca): an efficient approach for feature fusion of finger biometrics. *Multimed Tools Appl* 74(13):4469–4486
44. Yang J, Zhang X (2012) Feature-level fusion of fingerprint and finger-vein for personal identification. *Pattern Recogn Lett* 33(5):623–628
45. Yang J, Zhong Z, Jia G, Li Y (2016) Spatial circular granulation method based on multimodal finger feature. *J Electr Comput Eng*
46. Yang Y, Lin K, Han F, Zhang Z (2012) Dynamic weighting for effective fusion of fingerprint and finger vein. *PICA: Prog Intell Comput Appl* 1(1):50–61
47. Kumar A, Zhou Y (2012) Human identification using finger images. *IEEE Trans Image Proc* 21(4):2228–2244
48. Yang W, Huang X, Liao Q (2012) Fusion of finger vein and finger dorsal texture for personal identification based on comparative competitive coding. In: *2012 19th IEEE international conference on image processing (ICIP)*. IEEE, pp 1141–1144
49. Yang W, Huang X, Zhou F, Liao Q (2014) Comparative competitive coding for personal identification by using finger vein and finger dorsal texture fusion. *Inf Sci* 268:20–32
50. Asaari MAM, Suandi SA, Rosdi BA (2014) Fusion of band limited phase only correlation and width centroid contour distance for finger based biometrics. *Expert Syst Appl* 41(7):3367–3382

51. Kang BJ, Park KR (2010) Multimodal biometric method based on vein and geometry of a single finger. *IET Comput Vis* 4(3):209–217
52. Xiaoming X, Yilong Y, Gongping Y, Xianjing M (2013) Personalized fusion method based on finger vein and finger contour. *J Comput Res Dev* 9:015
53. Lu Y, Xie SJ, Yoon S, Yang J, Park DS (2013) Robust finger vein ROI localization based on flexible segmentation. *Sensors* 13(11):14339–14366
54. Zuiderveld K (1994) Contrast limited adaptive histogram equalization. In: Heckbert PS (ed) *Graphics gems IV*. Morgan Kaufmann, pp 474–485
55. Zhao J, Tian H, Xu W, Li X (2009) A new approach to hand vein image enhancement. In: *Second international conference on intelligent computation technology and automation, 2009. ICICTA'09*, vol 1. IEEE, pp 499–501
56. Zhang J, Yang J (2009) Finger-vein image enhancement based on combination of gray-level grouping and circular gabor filter. In: *International conference on information engineering and computer science, 2009. ICIECS*. IEEE, pp 1–4
57. Choi JH, Song W, Kim T, Lee SR, Kim HC (2009) Finger vein extraction using gradient normalization and principal curvature. *Proc SPIE* 7251:1–9
58. Lowe DG (1999) Object recognition from local scale-invariant features. In: *Proceedings of the seventh IEEE international conference on computer vision (CVPR'99)*, vol 2. IEEE, pp 1150–1157
59. Yang J, Yang J (2009) Multi-channel gabor filter design for finger-vein image enhancement. In: *Fifth international conference on image and graphics, 2009. ICIG'09*. IEEE, pp 87–91
60. Rohr K, Fornefett M, Stiehl HS (1999) Approximating thin-plate splines for elastic registration: Integration of landmark errors and orientation attributes. In: *Biennial international conference on information processing in medical imaging*. Springer, pp 252–265
61. Maio D, Maltoni D, Cappelli R, Wayman JL, Jain AK (2004) FVC2004: third fingerprint verification competition. In: *ICBA*, vol 3072. Springer, pp 1–7
62. Zhang L, Zhang L, Zhang D (2009) Finger-knuckle-print: a new biometric identifier. In: *2009 16th IEEE international conference on image processing (ICIP)*. IEEE, pp 1981–1984
63. Yang W, Yu X, Liao Q (2009) Personal authentication using finger vein pattern and finger-dorsa texture fusion. In: *Proceedings of the 17th ACM international conference on multimedia*. ACM, pp 905–908
64. Hofbauer H, Uhl A (2016) Calculating a boundary for the significance from the equal-error rate. In: *Proceedings of the 9th IAPR/IEEE international conference on biometrics (ICB'16)*, pp 1–4
65. Prommegger B, Kauba C, Uhl A (2018) Longitudinal finger rotation—problems and effects in finger-vein recognition. In: *Proceedings of the international conference of the biometrics special interest group (BIOSIG'18)*, Darmstadt, Germany

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

